University of Namur

THESIS / THÈSE

DOCTOR OF ECONOMICS AND BUSINESS MANAGEMENT

Distinctive Strategies in the Equity Mutual Funds Industry

Vanhomwegen, Henri

Award date:
2020

Awarding institution:
University of Namur

UNIVERSITÉ DE NAMUR

Faculty of Economics, Social Sciences and Business Administration
Center for Research in Finance and Management (CeReFiM)
Namur Institute for Complex Systems (naXys)

Distinctive Strategies in the Equity Mutual Funds Industry

Henri Vanhomwegen

A Thesis Submitted in Fulfillment of the Requirements
for the Degree of
Doctor of Economics and Business Management

Jury President

Professor Stéphane Faulkner, Université de Namur

Supervisors

Professor Jean-Yves Gnabo, Université de Namur

Professor Oscar Bernal Diaz, Université de Namur

Jury members

Professor Sophie Béreau, Université de Namur

Professor Christophe Boucher, Université de Paris Nanterre

Professor Alexandre Girard, Université Saint-Louis

Professor Mikael Petitjean, Université Catholique de Louvain, IESEG School of Management (Lille)

Defended on September 2^{nd} 2020

Rue Grandgagnage 19 B - 5000 Namur (Belgique)

 Imprimé en Belgique
ISBN : 978-2-39029-116-9
Dépôt légal: D2020188115

Acknowledgements

This PhD thesis is the pinnacle of the last four years of my life, which in retrospect feel like they passed in a heartbeat. To be completely honest, sometimes during my PhD, it felt more like a life sentence than a heartbeat, but that is the beauty of this journey which makes you feel emotions on a spectrum wider than you could imagine possible. My doctorate has truly been the most intense and enriching experience I can recall, and I will always thank myself for taking that leap and applying for this position.
Yet, I would not be were I am today if it wasn't for the incredible support, trust and encouragements I received from numerous individuals. What follows is my attempts to thank them and pay tribute to their generosity.

I wish to start with whom I kindly call the "instigator of it all": Sophie Béreau. If you had not seen in me what I certainly could not see in myself and pushed me to strive for more, I am a 100% positive that I would not be writing a doctoral thesis acknowledgement section at this moment. Back when I was pursuing my master thesis under your supervision, your inexplicable faith in me resurrected my self-confidence and pushed me to reach for the moon, which ultimately led me to land this PhD. For that, I will be eternally grateful to you.

Now for the piece of resistance, I would also like to express my deepest gratitude to Jean-Yves Gnabo, my supervisor, for his mentoring, sometimes his patience, but most of all for his unwavering support. These last few years, you have done the best you could to keep me on track, motivating me when I was at my lowest and kindly bringing my back to earth when at my highest! Your unrivalled mastery of sarcasm, encyclopedic culture, infinite wisdom and our shared love of gastronomy, have made this entire experience not half as painful as I feared it would be.

I am also very grateful to my co-supervisor Oscar Bernal for his support and the time he devoted to this thesis. Furthermore, I would like to sincerely thank the members of my jury: Christophe Boucher, Stephane Faulkner, Alexandre Girard and Mikael Petitjean for their invaluable comments and especially for taking some time off of their busy schedule to evaluate my thesis. I would like to thank the entire administrative staff of the Unamur too, and more particularly Pierrette Noël for the incredible work she does and her kindness. I thank the BNP Paribas research chair in risk management and everyone involved in the project, especially Marnix Arickx for his insights on the asset management industry and Sarah Marteleur and Morgane Belin for their organisational skills. I am are grateful to Anne-Florence Allard, Zacharias Bragoudakis, Kris Boudt, Arnaud Dufays, Deniz Erdemlioglu, Georges Hubner and Tina Zumer for their helpful comments and suggestions. I would also like to acknowledge the participants at the 22^{nd} and 23^{rd} annual conference on macroeconomic analysis and international finance, the 16^{th} and 17^{th} econometric days in Paris Nanterre, the Belgium Financial Research Forum 2018 as well as the international Phd colloquium in HEC liège. The seminar participants at CRED workshop and Naxys seminar (Unamur), the Vrij universiteit brussels and the IESEG school of Management Lilles for their feedbacks and valuable discussions.

To my colleagues, meeting you was already enough to make this journey worth it. Joey Soudant, words cannot expressed how incredible, kind and brilliant you are. Sharing our office this last 4 years has been an absolute blast and I could not have hoped for a better person to live this experience with. We have been through thick and thin together but have always pulled through, so I wish you the best of luck for what's to come! To Auguste, Camille, Christian, Doux, Eli, François, FX, Lhorie, Ludovic, Marie S., Marie D., Modeste, Victor, Jérome, Nathalie, Pierre and all the others: Thanks you for our memorable times together, our after-works, Yayami sessions, Christmas parties and lunches "au bord de l'eau". I will cherish those memories and hope to make many more with you all.

To Link-Matthieu, having you by my side through this entire process has been crucial to my balance and hapiness, especially considering how bonkers the last few months have been. Your unconditional love and support have made me a better and stronger person and allowed me to never give-up in the face of adversity. Therefore, I would like to dedicate this thesis to you, but just keep in mind that you are not the only "doctor" anymore love.

Finally, to my family, you have always been the cornerstone of my happiness. Papa, Maman, Paul and Louis, I love you all very much and thank you from the bottom of my heart for having always supported me. Now, time to get a real job as you would say!

Contents

List of Tables

General Introduction

This thesis — as suggested by its title – aims to investigate whether distinctive strategies (*i.e.* strategies deviating from a norm) implemented by active managers[1] could provide added value to their investors. This debate has been at the heart of the financial literature since the 1950's and led to the current consensus which points to the inability of active managers to generate abnormal performance persistently. Yet, seminal papers on which this consensus is based were mostly conducted prior to the 2000's and exclusively focused on US data, arguing than the results were canon to the rest of the industry. Hence, we highlight two weaknesses of the current literature. First, in the last 20 years, the global asset management industry has witnessed drastic changes in its structure, its competition level and its size, fostering new studies actually finding evidence against the consensus. Second, the European asset management market has grown to be the second largest on a global scale (after the US) and recent studies uncovered that its properties were substantially different from the ones of the US market. Therefore, we argue that conclusions drawn by seminal papers regarding mutual funds performance and managerial skills across the globe may not hold in current market conditions. Against this background, our aim through this thesis is to contribute to this research body by studying how investors (either individual or institutional) around the world, and especially in Europe, could distinguish effective strategies and benefit from actively managed funds. Taking into account the increased heterogeneity and complexity of the industry, its growth and development worldwide, we argue that studying how investors may identify the best potential opportunities (i.e. best fund managers) relatively to those available has never been more crucial. This thesis is thus divided into three paper-based chapters, each proposing an original approach that contributes to this specific debate. Chapter 1 proposes a market based approach — usable by individual or institutional investors alike— to distinguish active managers implementing distinctive strategies in the European market. Chapter 2, using portfolio holdings data, investigates strategies using portfolio con-

[1] As opposed to passive managers which simply track a specific indexes

centration in stocks and/or risk exposures to generate abnormal returns in the European market[2]. Finally, chapter 3 looks at socially responsible investments (SRI) funds in the US and challenges the conventional wisdom which states that the restrictions imposed by ethical considerations on their investment universe will necessarily lead them to under-perform more conventional funds.

The Asset Management Industry Evolution

The asset management industry is an integral part of the global financial system, as the banking and insurance industries are. Yet, it stands out from its counterparts given its diversity and unique structure. Indeed, asset management firms do not own their assets under management (AuM)[3] but rather owe fiduciary duties to clients, hence have the obligation to act in their best interest. As such, clients' assets are not guaranteed[4], yet are not tied to the asset management firm status. In other words, the firm assets and its client assets are legally separated and the later cannot be used in case of a failure.

Turning to the industry's diversity, on a global scale the number of investment strategies offered through the different types of investment funds such as —to cite a few– mutual funds (equity, bonds, mixed or money market), index funds, hedge funds and private funds is staggering. Yet, overall the strategies can be distributed into two main management styles: the so-called *passive* and *active* ones. A passively managed investment fund proposes to track a specific index as closely as possible (buy and hold strategy), while an actively managed funds will supposedly follow an original idea designed to outperform a referential benchmark or the overall market. The underlying motivations behind investing in a passive strategies are that it is cheaper (no fees linked to managerial skill, high turnover or screening and monitoring) and less complex (it simply tracks a well defined market index). However, the potential returns are low as a passive fund will never outperform the benchmark it tracks and the funds' managers are bound by the index, and as such have no investment flexibility. On the other hand, active strategies can be attractive as they are designed to beat their benchmark, and thus offer more potential returns. Moreover, they provide a greater flexibility to the managers given that they are not limited to an index. Yet, the potential downfalls in returns are greater than with index strategies (riskier) and the costs of operating are usually much greater which may eclipse the added value generated by a manager. In recent years, new managerial styles arose, some trying to blur the frontier between active and passive

[2]Given the difficulty and cost of gathering and consolidating data from portfolio holdings, we argue that this methodology is better suited to institutional investors

[3]As such these assets are not part of their balance sheets

[4]They are no guarantee that any investment amount will be recovered, in other words managers do not back-stop investment losses.

management, others trying to incorporate ethical criteria into the investment decisions. For instance, smart beta funds aim to beat a referential benchmark by variating the weights of its underlying components, thus passively tracking the index (as no changes in compositions is operated) while actively choosing the optimal weights for each positions. Socially responsible investment (SRI) funds on the other hands, aim to uphold the values of sustainable finance, notably by evaluating investment opportunities alongside three axes, environmental, social & governance (ESG) and as such go beyond the traditional paradigm of risk vs returns. Moreover, the industry's diversity was even further exacerbated by recent changes witnessed in its structure and size. Indeed, the global asset management has grown by approximately 140% between 2002 to 2017, going from managing approximately \$32 trillion (BCG 2010) to more than \$77 trillion[5] (BCG 2019), from which \$68 trillion are managed by mutual funds and \$9 trillion by alternative funds such as hedge and private funds. More strikingly, that growth was partly lead by the rise in popularity of passive strategies, notably by the introduction of a new type of fund: the exchange traded funds (ETFs). ETFs are index funds which have the particularity to be publicly offered on stock exchanges, these funds have been attracting substantial interests in the market during the last decade as they grew from approximately managing \$417 billion in 2005 to \$4.4 trillion by the end of 2017 (EY 2017). This rise in popularity of passive strategies has effectively changed the structure of the market as it went from representing 9% (\$3 tn) of all AuM in 2003 to more than 19%(\$14 tn) in 2017, while active management went from 82% (\$26 tn) to 75% (50 tn) BCG (2019). This is quite a significant shift as active management was once the only available option for investor willing to delegate their wealth management Cremers, Fulkerson & Riley (2019). Indeed, passive management is a relatively newer option[6] introduced by the first ever index tracking funds: the Vanguard 500 in late 1976 Morningstar (2011). As stated above, the management fees of such funds are usually much lower than their active counterparts, their growth has thus increased the competitive pressure on active funds, and compelled them to lower their fees in order to remain attractive. According to Cremers et al. (2019), it went from 1.06% in 2000 to 0.78% in 2017.

Mutual Funds Performance: From Jensen To Carhart

From an academic point of view, the sheer size of the current active asset management industry remains puzzling. Indeed, over the last 50 years of research —since Jensen (1969)

[5]As a comparison the banking and insurance industries represented \$124 (Deloitte 2019) and \$32.8 (FSB 2019) tn respectively at the same period.

[6]Although the concept of market portfolio was already promoted by both academics and practitioners for years before Morningstar (2011)

seminal work on mutual-funds performance— the general consensus in the literature points to the average inability of active managers to cover their costs and persistently create value for their investors, therefore finding no evidence of persistent managerial skills[7]. Management skills has undoubtedly been one — if not the most— topical subject in the financial literature. The vast majority of studies defines skills as the ability to create after fees abnormal returns (*i.e* in excess of a referential benchmark), which is traditionally referred to as the "net alpha". Computing such net alpha sprung the creation of several factor models aiming to best evaluate the actual abnormal performance generated by each manager. The first model, devised by Jensen (1968), is a single-factor models based on the capital asset pricing model (CAPM) by Sharpe (1964), Lintner (1965), Mossin (1966) & Treynor (1961). The model measures the performance of a portfolio with respect to the one predicted by the CAPM, as such it can be understood as the portfolio excess-return with respect to a referential benchmark (the world portfolio or SP500 for instance). Unfortunately, this model suffers from the same limitations as the CAPM, which are the very strong assumptions on which it builds. For instance the CAPM assumes full market efficiency and rationality from the market participants. In order words, it assumes that their is no information asymmetry in the market, that prices reflect at every moment all information available, and that investors based on these information will make the most optimal choice possible. As such, it falls prey to the joint testing hypothesis problem. The first hypothesis, when testing for abnormal returns, is that the CAPM is correctly specified, the second that market are fully efficient. Therefore, when abnormal returns are observed it is impossible to know whether the fund actually managed to generate abnormal returns due to its manager's skills or if it is due to market inefficiencies which the models does not control for (i.e the model is incorrectly specified). Indeed, long lived market inefficiencies have been documented in the literature, compromising the reliability of the CAPM results, given that if agent were truly rational and had access to the same information, they would try to take advantage of such price anomalies which would ultimately cancel them out. One of the first anomaly detailed in the literature (see, Banz 1981) was the historical outperformance of small capitalisation with regards to large ones, labelled the size premium. An other anomaly pertains to the historical outperformance of value stocks with regards to growth stocks[8]. The literature as used multiple ratios to segregate growth from value stocks and asses their relative performance, with notably the price earning ratio (PER) used by Basu (1977) or the price-to-book (P/B) ratio used by Fama & French (1993). These observations led Fama & French (1993) to introduce a multi-factor model building on

[7] The most notable studies include: Sharpe (1991), Fama & French (1993), Carhart (1997),Fama & French (2010)

[8] Value stocks are those considered to be traded under their fundamental value while growth stocks are those traded above

the same premise as the CAPM, yet attempting to take into account these specific market anomalies which could explain performance and be mistaken as signs of managerial skills. They authors introduced the 3 factor model which not only takes into account the market premium (RMRF) but also the historic outperformance of value stocks with regards to growth stocks (HML) and of small stocks with respect to large stocks (SMB), to assess the managerial skills and thus try to mitigate the joint testing hypothesis by consolidating the model. Then, building on Jegadeesh & Titman (1993) which observed that past winner (looser) stocks tended to follow short term upward (downward) trend in their performance[9] Carhart (1997) added a fourth factor to the former model to control for momentum (MOM), thus creating the four factor model. These models have enjoyed an incredible popularity in the financial literature given that they are now what seems to be the *de facto* choice to measure the performance of financial instruments (in the equity market). Other models have also gained popularity[10] over the years, with notably Daniel, Grinblatt, Titman & Wermers (1997*a*) which introduced an alternative to factor model by developing a holding based measure of performance that creates a benchmark based on the fund's portfolio constituents and then evaluates the difference in returns between the funds and its artificial benchmark, or Hoberg, Kumar & Prabhala (2017) which introduced a measure of funds performance in excess of their peers.

Active Management Versus Passive Management: The Debate Goes On

As described before, the industry as evolved and conclusions drawn by seminal studies may not hold in the current setting. This is notably highlighted by numerous recent academic studies finding new evidence against the general consensus, and pointing to the ability of some skilled managers to create value for their investors (see Cremers et al. 2019 for a review). Studies regarding portfolio concentration (see for instance Kacperczyk, Sialm & Zheng 2005, Cremers & Petajisto 2009, Amihud & Goyenko 2013, Choi, Fedenia, Skiba & Sokolyk 2017) find evidence that managers concentrating their assets where they have informational advantages are able to generate positive abnormal returns, thus showing signs of persistent managerial skills. Others studies such as Sun, Wang & Zheng (2012), Vozlyublennaia & Wu (2017) and Hoberg et al. (2017) find that in general funds implementing strategies deviating from their peers tends to out-perform them. Thus, the debate regarding managerial skills and how investors could distinguish it is still very much alive. Moreover, to this day a massive gap

[9]Therefore that strategies buying past winner stocks and selling past looser stocks had historically out-perform the market.

[10]Although not as much as the ones of Fama & French (1993) nor Carhart (1997)

in the literature concerning funds performance remains. Indeed, the vast majority of studies on the asset management are conducted using US data and draw conclusions supposedly translatable to all markets. This is problematic given the fact that the European Union is the second biggest market worldwide regarding asset management[11], yet more importantly that it does not have the same characteristics as its US counterpart. For instance, in what to this date represents the most comprehensive study on the asset management industry, Ferreira, Keswani, Miguel & Ramos (2013) found evidence that the diseconomies of scale, largely documented for US funds, were not observed in the EU market where the author found actual evidence of economies of scale[12]. The overall structure of the market is also quite different as the EU implemented general laws encompassing the asset management industry for all member states (UCITS V[13] for instance), but let each member the liberty to implement specific fiscal laws which could foster the domiciliation of certain types of funds[14]. Thus, As noted by Banegas, Gillen, Timmermann & Wermers (2013), improving our understanding of European funds and their strategies is critical for a large set of actors including investors, regulators, and policymakers.

Structure Of The Thesis

As detailed before, the thesis is divided into three paper-based chapters revolving around a common aim which is to provide investors (be them individual or institutional) with tools to identifies successful strategies from active managers.

Chapter 1 is entitled "Making a difference: European mutual funds distinctiveness and peers' performance" and was written under the joint supervision of Prof. Sophie Béreau and Prof. Jean-Yves Gnabo, my PhD supervisor. In this study, the sample consist of 4,957 funds over the 1999-2016 period. The most comparable databases are used by Banegas et al. (2013) and Graef, Vogt, Vonhoff & Weigert (2018), who also cover the European segment of the mutual funds industry. The former uses 4,200 funds from 1988 to 2008 at a monthly frequency while the later focuses on 1,464 European funds from 2001 to 2017 at a semi-annual frequency. We can further add to these studies the research by Ferreira et al. (2013), which provides a worldwide analysis of mutual funds with 4,438 European funds out of the

[11]The EU market represented 37% of the global asset management industry by the end of 2016, while the US accounted for 51% (EFAMA 2017)

[12]This is notably explained by the fact that on average EU funds are much smaller than their US counterpart.

[13]https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32014L0091&from=FR

[14]Luxembourg for instance is the highest hub worldwide of global and international funds Lang & Köhler (2011), because of its attractive fiscal laws EY (2020)

12,577 considered between 1997 to 2007. Given that our dataset is more extensive and span on a longer time horizon we argue that we provide additional insights into the European equity mutual funds industry. More specifically, in this chapter we evaluate if "distinctive" strategies allow funds to outperform their peers and test this hypothesis in different market conditions. Our main intuition is that funds' competition primarily takes place within styles (*e.g.*, Eurozone Large-Cap, France equity). Yet, that given the broad definition of these styles, it leaves funds with the possibility to implement innovative strategies in order to "stand-out from the crowd" (Vozlyublennaia & Wu 2017) and outperform their closest competitors. This study principally builds on the ones of Brown & Goetzmann (1997) and Sun et al. (2012) which find that funds which correlate less with their core category tend to outperform. However, we depart from those studies in the way we uncover each fund's peers and compute their relative performance. To do so, we rely on the adaptive forgetting factor for evolutionary clustering and tracking (AFFECT) cluster algorithm devised by Xu, Kliger & Hero Iii (2014). Its advantages are twofold, (i) it allows to retrieve the funds peers endogenously through time, thereby addressing changes in the number of styles, funds potential shifts in style, and the entry and exit of funds, and (ii) it does so by only relying on market data, thus avoiding to use portfolios holdings ones which can be tedious to obtain and consolidate. Equipped with our endogenous categories, we first apply the strategy distinctiveness index (SDI) devised by Sun et al. (2012), which measures the strength with which a fund distinguish itself from its style. Secondly, we compute the relative performance of funds with respect to their peers by relying on the out-performance ratio from Ardia & Boudt (2018), which takes care of the "false discovery" issue detailed by Barras, Scaillet & Wermers (2010). Our main results highlight a strong, robust, and positive impact of strategy distinctiveness on financial performance. Yet, the effect is non-linear as its impact decreases as a fund becomes too distinct.

I have presented this chapter at multiple national and international conferences and seminars. Notably, the 22^{nd} International Conference on Macroeconomic analysis and International Finance (ICMAIF) at the University of Cretes (2018) , the Belgium Financial Research Forum at the National Bank of Belgium (2018), the 17^{th} "Journée de l'économétrie" in Paris Nanterre (2018), and the Vrij Universiteit brussels (VuB) and IESEG School of Management (Lille) internal seminars (2018 and 2019 respectively). Finally, in September 2019, this chapter was accepted for publication in the international peer-reviewed journal "Finance". Finance is ranked 2 in the CNRS categorization of Journals in Economics and Management.

Chapter 2 is entitled "Portfolio concentration and financial performance: insights from domestic and global equity mutual funds" and was written under the supervision of Prof.

Jean-Yves Gnabo. In this paper, we use information on portfolio holdings and the sample is made of 1,746 funds over the 2003Q1 to 2016Q4 period. To the best of our knowledge this is one of the most extensive and comprehensive database considering the portfolio holdings of European equity mutual funds. Indeed, information on holdings is not easy to access and is difficult to consolidate. Funds holdings are not publicly available at any time as they are a testament of one's strategy, however funds have — by law — to disclose their holdings twice a years (in some countries 4 times a year), then data provider are able to collect and store such disclosures (of potentially thousands of funds) for their clients to have access to. Yet, a single fund in its lifetime may have invested in several thousands positions on which information should be retrieved and consolidated in order to conduct the study. Hence, this motivated the use of Morningstar Direct throughout this thesis, given its comprehensive coverage across countries for funds and securities, as well as the availability of information regarding the funds' complete historical portfolio holdings. Closest studies in terms of dataset are the ones of Graef, Vogt, Vonhoff & Weigert (2019) and Franck & Kerl (2013). The former consider 1464 European funds with portfolio holdings at a quarterly frequency over the 2001-2017 period, the later collected portfolio holdings for 4315 European funds yet restricted their sample to the 20005-2009 period at a semiannual frequency. The main motivation behind this study pertains to recent literature developments pointing to the ability of concentrated funds to significantly outperform their benchmark (Brands, Brown & Gallagher 2005, Cremers & Petajisto 2009) and — in some cases— the market (Kacperczyk et al. 2005, Amihud & Goyenko 2013, Choi et al. 2017). Therefore, in this chapter, we evaluate if concentrated strategies are indeed reflective of higher managerial skills and enable to reap positive risk-adjusted returns. We build on Cremers & Petajisto (2009) and define concentration along two axes: stock selection (overweighting of core positions) and risk factor exposures (effective number of risk factor bets). We propose to do so by addressing several caveats traditionally overlooked by the literature. The first caveat concerns measures of concentration defined against a referential benchmark[15], as these measures are by design extremely sensitive to the chosen benchmark. As demonstrated by Amihud & Goyenko (2013), if a fund invests in multiple styles, concentration estimations retrieved from benchmark-centric measures may be misleading.

The second caveat relates to the definition of the risk factors themselves. Indeed, the vast majority of studies are based on US funds with domestic strategies and as such rely on "trivial" economic-based US factors such as industry factors (Kacperczyk et al. 2005) or style factors[16] (Amihud & Goyenko 2013). However, when considering funds with highly

[15] Such as in Cremers & Petajisto (2009)

[16] Small cap premium (SMB), value premium (HML), momentum premium (MOM)

heterogeneous strategies (*e.g.*, global, international or domestic funds) — such as the case in Europe— their potential number becomes very large. Therefore, the literature on international finance proposes to group the factors into different factor models: country, industry and style[17]. Yet, it overlooks the possibility that factors from different models may proxy the same underlying risk (Lessard 1974, Huij & Derwall 2011) and therefore may lead to understate the actual risk concentration of mutual funds. To circumvent this issue, we propose to rely on a recent development of the risk budgeting literature which uses principal component analysis (PCA) to uncover endogenous and uncorrelated risk factors directly from the assets composing the portfolio. Then, we compute the effective number of factors to which each fund is exposed to, in order to evaluate the breadth of the underlying strategy (Grinold & Kahn 2000). Our intuition is that managers may focus their strategy on segments of the market in which they have some expertise, or conversely manage their risk exposures by actively spreading it through multiple factors. Relying exclusively on portfolio holdings data to compute the stocks concentration and risk factor exposures allows to avoids entirely the two aforementioned shortcomings.

Finally, crossing our two measures, we highlight that managers able to concentrate their holdings in few core positions, while still spreading their risk exposures through multiple risk factors, generate positive risk-adjusted returns and outperform their competitors, even in time of financial turmoil.

I presented this paper at the 23^{rd} ICMAIF at the university of Crete (2019), as well as at internal seminars at the University of Namur (2019).

Chapter 3 is entitled "SRI mutual funds' performance and investment universe" and was written under the supervision of Prof. Jean-Yves Gnabo. This study tackles the topical issue of socially responsible investment (SRI) funds' performance. Given that the literature on sustainable funds is more recent, and our wish to ground our study on existing ones to offer a different point of view on what one could expect from their performance, we decided to focus on US funds. Our dataset is made of 2,039 US portfolio for which holdings information as well as the funds' sustainability score are retrieved from 2002Q3 to 2018Q4[18]. SRIs managers aim to maximize financial returns, while constraining their investments to firms deemed as socially responsible, thus superimposing an ethical consideration to portfolio selection. According to traditional portfolio theory, such a constraint will mechanically worsen portfolio diversification —as it decreases a fund's number of eligible assets—, thus impacting its efficiency (*e.g.*, Markowitz 1952, Barnett & Salomon 2006, Renneboog, Ter Horst &

[17]See Huij & Derwall (2011) for a review on the subject

[18]the closest study to ours from El Ghoul & Karoui (2017), is made up of 2,168 funds at a yearly frequency from 2003 to 2015

Zhang 2008*b*). We challenge this narrative based on the fact that most managers already operate on a restricted list of assets determined by their style(s) or reference benchmark(s). Therefore, we argue that the potential negative impact on performance induced by an SRI strategy[19] may be alleviated by following a style (or multiple ones) giving access to numerous investment opportunities (large investment universe).

To test our hypothesis, we propose a novel measure of a fund' investment universe size (*IU Score*). To do so, we use Morningstar styles (e.g., Large-cap value, Small-cap blend) and their associated benchmark's holdings, which effectively reflect the entire US equity market. Then, using a portfolio overlap measure[20] —based on the Manhattan distance— we compute the similarity of each fund's portfolio to all available benchmarks. At each period, we measure the fund's *IU Score* as the total number of stocks held by the associated styles' benchmark scaled by the magnitude of their overlap with the focal fund. Finally, we cross Morningstar's *Sustainability Score* — which measures the SRI strength of each fund— with the *IU Score* to test whether SRI funds associated with larger universes fare better than their counterparts associated with smaller universes. Furthermore, we also test for their performance relative to more conventional funds associated with either large or small universes. Our main results highlight that there is no significant difference between SRI funds and conventional funds' performance level. However, when their performance is evaluated conditionally to their investment universe, we show that SRI mutual funds with the smallest universes consistently under-perform other funds in the market. We conclude that styles selection is critical for SRI managers in order to minimize the potential negative side effect of restrictions imposed by their ethical goals.

The paper is currently in submission for the 24^{th} ICMAIF at the university of Crete, and to the 3^{rd} annual conference of the Global Research Alliance for Sustainable Finance and Investment (GRASFI) at the University of Columbia (NY).

[19] See Renneboog, Ter Horst & Zhang (2008*a*), El Ghoul & Karoui (2017)

[20] Inspired by Cremers & Petajisto (2009) *Active Share* measure

Chapter 1

Making a Difference: European Mutual Funds Distinctiveness and Peers' Performance

1.1 Introduction

Whether an investment fund can consistently outperform its competitors is still highly debated in the academic literature. In their chase for higher "alpha", funds considered to be endowed with sufficient skills can decide to stand out "from the crowd"—to quote Vozlyublennaia & Wu (2017)—by developing distinct strategies. Meanwhile, others limit themselves to following the average behavior of their peers' competitors.

The main issue we explore in this study is whether it is worth differentiating oneself from one's peers in the mutual funds industry. We do so by addressing the two following questions: Do distinctive strategies enable a fund to outperform its peer competitors and what are the key drivers of distinctive strategies? Both questions are empirically examined with care in the poorly documented context of the European mutual funds industry. Our main findings are as follows. First, we demonstrate that European equity mutual funds (EEMFs) generate on average higher risk-adjusted performance than their close competitors when they follow "distinctive" strategies; however, this effect is non-linear. Hence, the marginal effect tends to decrease with the level of distinctiveness. In addition, consistent with the notion of "migration risk", we show that the transition toward more distinctive strategies can be costly, as a fast shift is associated with lower returns. Second, looking at the underlying motivations, we find that most distinct funds are young and small.

Béreau, S., Gnabo, J. Y., Vanhomwegen, H. (2020). Making a difference: European mutual funds distinctiveness and peers' performance. *Finance*, *41*(2), 7-51

These results provide new empirical evidence on funds' performance, adding to an already rich literature (see as general reference Ferreira et al. 2013 and the work by Sun et al. 2012 and Vozlyublennaia & Wu 2017 on style, distinctiveness, and fund performance). Specifically, we contribute to the line of research testing the ability of particularly "skilled" funds that implement innovative strategies to beat their competitors. The rationale behind this test can be summarized as follows. Competition among funds exists primarily within styles (*e.g.*, Eurozone Large-Cap Equity or France Large-Cap Equity), with funds developing investment strategies to outperform their style peers (Hoberg et al. 2017). As these styles are only broadly defined, they leave funds with sufficient latitude to differentiate themselves from direct competitors—other funds following the same style—to generate higher gains (see Daniel, Grinblatt, Titman & Wermers 1997*b*; Sun et al. 2012; and more recently Hoberg et al. 2017). Whether such an objective is achieved in practice in the asset management industry remains, however, an open question that academic research can help answer.

This question requires addressing several empirical caveats. Most notably, we need to identify the set of relevant competitors (*i.e.*, funds following the same style) with which each fund should be compared and determine a methodology to confront their respective performance. The first caveat stems from the lack of reliable information about fund style. As discussed by Sensoy (2009), for instance, self-reported styles are subject to strategic manipulation by investment funds (see also Hoberg et al. 2017), which casts doubt on their accuracy. To circumvent this issue, the literature proposes applying statistical tools to retrieve the set of institutions following the same style from the dependence between funds' characteristics such as returns on total net assets (TNA) (see Brown & Goetzmann 1997, Sun et al. 2012).[1] Another caveat concerns the procedure used to compare funds' performance with its competitors. Because typical "absolute" indicators of risk-adjusted performance such as three- and four-factor alphas are estimated quantities, they are not directly comparable across funds and require specific econometric treatment to account for sampling errors (Ardia & Boudt 2018). For instance, two funds can display different estimated alphas, but such a difference could be too small for the true unobserved alphas to be actually deemed as different. As subsequently described, each of these problems is far from trivial and needs to be addressed carefully to draw sound conclusions. In this study, we apply recent econometric methods as well as machine learning-type approaches to address these issues. In particular, we apply an endogenous dynamic clustering approach, the Adaptive Forgetting Factor for Evolutionary Clustering and Tracking (AFFECT) methodology proposed by Xu et al. 2014, to retrieve the number of styles in each period and allocate each fund to its style. We then compute the

[1] An alternative approach adopted by Cremers & Petajisto (2009) among others consists of using the similarity of portfolio holdings.

strategy distinctiveness index (SDI) of Sun et al. (2012). This metric enables us to set the distance between each fund and its style peers. Next, we implement the peer performance test developed by Ardia & Boudt (2018), which is immune to the so-called "false discovery" problem (see Barras et al. 2010) to assess the proportion of competitors beaten by a specific fund. Finally, we use panel regression analyses to both test the impact of the SDI on peer performance and reveal the determinants of the SDI. For the sake of comparison, we also propose alternative approaches for estimating our relationships of interest, including, for instance, either a specification with the original four-factor alpha (Carhart 1997) as a dependent variable to measure mutual fund performance or a specification with variables based on a clustering method applied to fixed length rolling windows as in Sun et al. (2012). Another obstacle to overcome in this research is accessing comprehensive data on European mutual funds. As noted above, compared with the U.S. mutual funds industry, the European market has been poorly described so far despite substantial growth over recent decades. To conduct the analysis, we thus collated a novel database comprising 4,957 EEMFs and totalling €1,014 trillion of capital under management by the end of 2016.

The closest study to ours is that of Sun et al. (2012). We build and extend their initial analysis in a number of ways. *First*, we focus on the mutual funds industry as opposed to the hedge funds industry in the reference study. The literature on the hedge funds and mutual funds industries has grown hand-in-hand over several decades, with contributions exploring separate as well as common issues. While hedge funds have long been considered to be far riskier and more aggressive than mutual funds, the growing risk taking in the mutual funds industry, as reported by Choi et al. (2017), makes the assessment of their strategy along with their resulting performance increasingly important. *Second*, we use data on European domiciled funds as opposed to U.S. domiciled funds in Sun et al. (2012). The mutual funds industry has been substantially growing outside the United States, with Europe being responsible for a large part of this growth. For example, between 1999 and 2016, the EEMF market expanded by more than 400% ($722 billion to $4.1 trillion) compared with 200% over the same period for the United States (ICI 2017). In total, the size of the EEMF market reached €3.8 trillion of assets under management (AUM) in 2016, representing 37% of that of the U.S. equity mutual fund industry and 23% of the global equity mutual fund industry (EFAMA 2017). Importantly, Ferreira et al. (2013) note—in their cross-country analysis of mutual funds—substantial differences in the determinants of fund performance in the United States and elsewhere in the world, casting doubt on whether U.S.-based findings can be extrapolated to other markets. They show, for instance, that the well-known relationship between fund size and performance differs markedly when comparing U.S. with non-U.S. domiciled funds. As noted by Banegas et al. (2013), improving our understanding of EEMFs, their environment, and their strategy is thus critical for a large set of actors including investors, regulators, and policymakers. To the best of our knowledge, evidence on the link between strategy distinctiveness and financial performance for European institutions is lacking. *Third*, we propose a modified version of Sun et al. (2012)'s procedure to compute the SDI. Such a procedure is based on a fully dynamic and endogenous clustering approach (hereafter referred to as an adaptive approach) as opposed to a simple rolling window approach (non-adaptive approach) in its original form. As discussed in Section 3, style identification is a central ingredient in the measure of strategy distinctiveness. Our approach should therefore provide a more accurate measure. *Fourth*, we adopt a more consistent approach to measuring funds' performance. Sun et al. (2012) define the competitive environment of each fund as the set of funds following the same style. Their strategy distinctiveness measure is consistently computed with respect to this specific set of competitors. However, in their analysis, a fund's performance is based on a risk-adjusted alpha, supposing implicitly that each fund competes with the whole universe of funds. In the spirit of Cohen, Coval & Pástor (2005) and other contributions on "peer alpha" (Hunter, Kandel, Kandel & Wermers 2014), we measure the fund's performance as a fund's risk-adjusted returns relative to the risk-adjusted returns of its competitors (*i.e.*, funds belonging to the same style). Then, we rely on statistical tests in our comparison to control for sampling noise.[2] *Fifth*, we explore the existence of non-linearity in the relationship between strategy distinctiveness and financial performance to assess the existence of an optimal distance or a transition cost when shifting toward an innovative strategy due to migration risk.[3]

Overall, our work provides new empirical evidence on mutual funds and the role of distinctive strategies. Our results support the existence of a strong and positive link between strategy distinctiveness and relative performance measured by the percentage of peers outperformed. This finding is robust to a series of sensitivity tests, including the use of traditional four-factor alpha to measure absolute performance. In this case, our findings mean that singular strategies improve performance above both (style) peer competitors and the rest of the market. Turning to the determinants of a fund's distinctiveness, fund age, volatility, and TNA values are found to be negative drivers of strategy distinctiveness. These results confirm previous results found for other segments of the asset management industry and other markets. In particular, our results are consistent with those of Sun et al. (2012) and Vozlyublennaia & Wu (2017), who, with the exception of age for the latter, find the same type of relationship

[2] For comparison purposes, we also provide regressions with the traditional four-factor model Carhart (1997) to measure risk-adjusted performance.

[3] Migration risk stems from the lack of knowledge or expertise of fund managers when they start implementing new strategies. It is usually associated with changes in style. However, this principle can arguably be extended to changes occurring within a style, as studied in our analysis when a fund shifts to more singular and innovative strategies.

between the aforementioned fund characteristics and strategy distinctiveness for U.S. hedge funds and U.S. mutual funds, respectively. We also unveil new features such as the impact of cluster-related characteristics on future performance. The results of the non-linear specifications provide interesting nuances for the analysis. Specifically, we highlight (i) the stronger effect of strategy distinctiveness on financial outperformance during the crisis period, (ii) the lower performance when the shift toward more innovative strategies is fast, and (iii) the existence of a tipping point beyond which the positive marginal effect of the strategy distinctiveness variable dies out. As a result, although following innovative strategies pays off on average, the effect is non-linear and exhibits a threshold level over which being too different becomes useless.

The remainder of this paper is structured as follows. In Section 2, we discuss our data. In Section 3, we describe our methodology used to dynamically assess funds' clusters, which then allows us to measure funds' distinctiveness among their peers and quantify how this characteristic affects performance and other aggregates over time. In Section 4, we review the EEMF industry and present the results from the regression analyses. Finally, Section 5 concludes.

1.2 Data

We create an original database on EEMFs. To that aim, we take advantage of micro-level data (*i.e.*, fund-level information) on equity mutual funds domiciled in a broad range of European countries. The EEMF data are extracted from the Morningstar Direct database. Morningstar is widely used in the asset management literature (see Sensoy 2009, Patel & Sarkissian 2017). Among the key advantages of the Morningstar Direct database are its comprehensive coverage across countries and over time as well as the presence of non-surviving and surviving funds, making it free from survivorship bias. Information on fund attributes is available at various frequencies. For the present study, we collect daily prices along with returns[4] and TNA values, both at a monthly frequency. We also retrieve several other fund characteristics such as fund age and flows (see Table 1.1 for the details).

[4]Morningstar does not adjust total returns for sales charges or redemption fees. However, the data account for management, administrative, and 12b-1 fees as well as other costs automatically deducted from fund assets.

Table 1.1: Variables of interest (1999-2016)

Variables	**Frequency**	**Units**	**Sources**
Return	Monthly	%	Morningstar
Price	Daily	Euro	Morningstar
TNA	Monthly	Billion Euro	Morningstar
Age	Monthly	Years	Morningstar
Flow	Monthly	Billion Euro	Morningstar
Assigned benchmark	Point		Morningstar
Vstocxx	Monthly		Macrobond

Note: Table 1.1 reports the original data frequency, unit and sources for our main variables of interest.

We recover 8,520 equity mutual funds for 1999–2016. Then, we apply successive filters consistent with the literature (Sun et al. 2012, Ferreira et al. 2013). For instance, we exclude funds with fewer than 10 observations and less than €10 million in AUM. We further exclude funds of funds, index tracking funds, and funds not traded in euros. In addition, styles representing less than 1% of the population of funds and associated funds are excluded. We end up with 4,957 funds dispatched into 22 distinct styles, as identified by Morningstar (see Table 1.2).

The TNA ($\sim$ €1 trillion) of selected funds by the end of 2016 amounted to almost 40% of our target industry (EFAMA 2017, Morningstar 2016). Table 1.3 lists the European countries and their corresponding populations of funds.

Compared with the existing literature, the strength of our database is threefold. First, it covers a large number of funds. As a point of comparison, the aforementioned studies of Sun et al. (2012) and Vozlyublennaia & Wu (2017) that include the determinants of investment fund performance in the U.S. market rely on 3,896 and 3,519 funds, respectively. Second, information is available at a relatively high frequency, with the key attributes observed every month. Third, we embrace nearly two decades of European data that cover both crisis and post-crisis periods. To the best of our knowledge, the most comparable databases are used

Table 1.2: Benchmark and styles (December 2016)

Styles	**Benchmarks**	**# of funds**	**€billion**
Europe Small-Cap Equity	MSCI Europe Small Cap NR EUR	82	17.47 €
Asia-Pacific incl Japan Equity	MSCI AC Asia Pacific NR USD	60	9.14 €
Sector Equity Consumer Goods & Services	Cat 50%MSCI Wld/CD NR&50%MSCI Wld/CS NR	46	14.15 €
Global Emerging Markets Equity	MSCI EM NR USD	228	114.55 €
Europe Large-Cap Value Equity	MSCI Europe Value NR EUR	89	26.34 €
Global Large-Cap Blend Equity	MSCI World NR USD	631	188.32 €
US Large-Cap Blend Equity	Russell 1000 TR USD	146	58.65 €
Europe Large-Cap Blend Equity	MSCI Europe NR EUR	519	160.65 €
Asia ex Japan Equity	MSCI AC Asia Ex Japan NR USD	84	29.79 €
US Large-Cap Growth Equity	Russell 1000 Growth TR USD	58	36.15 €
Eurozone Large-Cap Equity	MSCI EMU NR EUR	364	86.62 €
Global Large-Cap Growth Equity	MSCI World Growth NR USD	107	49.56 €
Global Large-Cap Value Equity	MSCI World Value NR USD	98	45.80 €
Europe Large-Cap Growth Equity	MSCI Europe Growth NR EUR	65	33.24 €
Japan Large-Cap Equity	Topix TR JPY	100	38.96 €
Sector Equity Technology	MSCI World/Information Tech NR USD	48	12.60 €
Sector Equity Healthcare	MSCI World/Health Care NR USD	57	19.05 €
France Large-Cap Equity	Euronext Paris CAC 40 NR EUR	100	23.83 €
Asia-Pacific ex-Japan Equity	MSCI AC Asia Pac Ex JPN NR USD	59	27.88 €
Emerging Europe Equity	MSCI EM Europe NR EUR	42	5.37 €
Italy Equity	MSCI Italy NR EUR	45	9.07 €
Spain Equity	MSCI Spain NR EUR	69	6.95 €
	Total	**3097**	**1,014.13 €**

Note: Table 1.2 lists the 22 styles identified by Morningstar Direct as well as their associated benchmark. The last two columns report the number of funds (in 2016Q4) belonging to each style along with their total net assets.

by Banegas et al. (2013) and Graef et al. (2018), who also cover the European segment of the mutual funds industry. We can further add to these studies the research by Ferreira et al. (2013), which provides a worldwide analysis of mutual funds. In their analysis, European domiciled mutual funds account for 4,438 of the 12,577 mutual funds in total and observations span from 1997 to 2007. Graef et al. (2018) include 1,464 European funds from 2001 to 2017 at a semi-annual frequency, whereas Banegas et al. (2013) use 4,200 funds from 1988 to 2008 at a monthly frequency. Therefore, we provide additional insights into the European equity mutual funds industry.

1.3 Empirical Methodology

The empirical analysis aims to investigate the impact of EEMFs' distinctiveness (relative to peers) on future performance as well as its causes. To this end, we propose a modified

Table 1.3: EEMFs Domiciles

Domicile	**# of funds**	**% of total funds**
Luxembourg	2074	42%
France	849	17%
Ireland	424	9%
Germany	312	6%
Spain	293	6%
Italy	256	5%
Belgium	144	3%
Austria	127	3%
Netherlands	126	3%
Finland	105	2%
United Kingdom	76	2%
Switzerland	42	1%
Portugal	28	1%
Slovenia	15	0%
Sweden	14	0%
Liechtenstein	11	0%
Norway	10	0%
Denmark	10	0%
Greece	10	0%
Guernsey	8	0%
Monaco	6	0%
Andorra	5	0%
Malta	4	0%
Jersey	3	0%
Isle of Man	1	0%
Gibraltar	1	0%
Estonia	1	0%
Poland	1	0%
Hungary	1	0%

Note: Table 1.3 reports the number of funds per country (*# of funds*) as well as their % share with respect to the total number of funds.

version of the approach developed by Sun et al. (2012) that addresses several of the econometric caveats associated with the data. In Section 3.1, we describe our empirical approach to retrieve endogenous styles and compute fund-level strategy distinctiveness over time. In Section 3.2, we present the regression setting.

1.3.1 SDI

To capture the degree of distinctiveness or, on the contrary, the degree of similarity among EEMF investments at both the system-wide and the individual levels, we follow the general setup proposed by Sun et al. (2012), which relies on a measure of the distance between funds' net asset returns. The underlying motivation for using such a proxy is simply that we expect closer net asset returns for funds exhibiting similar portfolios, which reflect a similar investment strategies. An alternative is to directly rely on information on the portfolio holdings of funds (Cremers & Petajisto 2009, Gupta-Mukherjee 2013) and measure the deviation from a passive benchmark. However, those data are not easily accessible[5] and net asset returns are available at a higher frequency, providing more flexibility to describe the evolution of commonalities over time.

As previously noted, our main measure is constructed by following the setup developed by Sun et al. (2012) to compute the SDI. In their work, the SDI assesses how distinct and unique a fund strategy is relative to its peers. Formally, the SDI measure is calculated for each fund i as follows:

$$\begin{aligned} SDI_{i,t} &= 1 - corr(r_{i,t}, \mu_{I,t}) \\ &= 1 - \frac{\sum_{t=1}^{24}(r_{i,t} - \bar{r}_i)(\mu_{I,t} - \bar{\mu_I})}{\sqrt{\sum_{t=1}^{24}(r_{i,t} - \bar{r}_i)^2 \sum_{t=1}^{24}(\mu_{I,t} - \bar{\mu_I})^2}} \\ \text{Where} \quad \mu_{I,t} &= \frac{\sum_{i \in I} r_{i,t}}{count(i \in I)} \end{aligned} \tag{1.1}$$

Thus, this metric corresponds to 1 minus the correlation between the fund's returns ($r_{\mathrm{i,t}}$) and the average returns of all funds in the same cluster or style indexed by I ($\mu_{I,t}$). The higher the SDI, the more distinct is the investment strategy of a fund with respect to its

[5]Working with portfolio holdings is generally costly because the data collection is long and taxing. It also often limits the number of funds to be considered as well as the data frequency. For instance, Cremers & Petajisto (2009), Gupta-Mukherjee (2013) include no more than 3000 mutual funds at a quarterly frequency.

peers from the same cluster. To obtain a time-varying measure consistent with the monthly frequency of our analysis, correlations are computed as realized correlations, namely within-month correlations from daily returns in our case.[6]

One critical aspect of the methodology lies in the identification of fund styles. A fund's style governs the broad direction of the investment strategy (*e.g.*, concentrating on a specific region or class of assets). Within a style, funds can depart from their peers by making decisions depending on their skills or innovativeness. Therefore, pairing funds according to their style is a pivotal step in subsequently gauging the innovativeness of each strategy relative to style-based peers. For fund managers, peer funds following a similar style are particularly relevant for comparison purposes because they constitute their natural "rivals" or "competitors"[7], as stated by Hoberg et al. (2017) (see also DiBartolomeo & Witkowski 1997, Brown & Goetzmann 1997). To allocate mutual funds into styles over time, we follow two strategies that both infer the style directly from the data. First, we rely on a traditional rolling window approach, which we refer to as "non-adaptive". Second, we apply an "adaptive" clustering method (Xu et al. 2014). In the forthcoming sections, we elaborate on both approaches.

1.3.1.1 Non-adaptive cluster algorithm

Ample anecdotal evidence exists in the press and more formal evidence exists in the academic literature that self-reporting of style by funds—the so-called prospectus—can provide a misleading picture of their actual style-based strategy. The simple reason is that a financial institution may have a strategic interest in misreporting its style by announcing ex post to have followed a style to mask poor performance under its true strategy (see Sensoy 2009, Cremers & Petajisto 2009, Hoberg et al. 2017). In response, several studies have proposed alternative approaches to endogenously retrieve style-based categories and the associated fund allocations from data. The identification procedure builds on the intuition that funds that are similar or close to each other should be placed in a common group, whereas those appearing more distant should be in different groups. Each group then stands for the unobserved underlying style. Equipped with distance measures for all pairs of funds, alternative clustering techniques can be applied to allocate them into consistent groups. In practice, the

[6] Sun et al. (2012) apply rolling windows of 24 months over the entire sample. Our approach of using within-month correlations avoids us making an arbitrary choice about the size of the rolling windows. It is also more consistent with the monthly frequency used in the regression analysis.

[7] Hoberg et al. (2017) consider institution-specific competitors. As discussed in Section 1, we adopt a more conventional definition of the set of competitors: all funds belonging to the same style are assumed to be direct competitors. This approach has the advantage of being transitive. In other words, if fund A is a rival to fund B and fund B a rival to fund C, then we assume that A and C are rivals, which is not guaranteed under Hoberg et al. (2017)'s methodology.

allocation of mutual funds into accurate similarity-based categories is challenging. Here, we follow Brown & Goetzmann (1997, 2001) and Sun et al. (2012) to group our sets of funds in a consistent manner by applying clustering algorithms on funds' net asset returns. The seminal approach on which we rely is the well-known "K-means" algorithm (Hartigan 1975), which has a simple and intuitive rationale. Hence, considering a set of n observations and K clusters specified ex ante[8], the "K-means" algorithm aims to find a partition such that the squared difference between the empirical mean of each cluster and the points in the cluster is minimized. In other words, the observations are grouped into K clusters such that each belongs to the group with which the associated mean exhibits the closest distance. The process is traditionally iterative. First, the K observations of the dataset are assigned as the initial cluster means. Second, observations are gathered according to their nearest mean, which is computed as the smallest Euclidean distance.[9] Then, the process iterates between step one and step two until the results converge to a final segmentation of the data. An important shortcoming of the previous algorithm concerns its static nature because it has been designed to address non-dynamic systems. In some cases, however, we observe time variations in both the composition and the number of clusters. The asset management industry is one such case. In essence, the industry evolves continuously over time given the appearance and disappearance of funds and variations in styles due to changing market conditions and perceived profitability. One way to address the time-varying nature of styles and associated clusters is to implement static methods over rolling windows, as in Brown & Goetzmann (1997, 2001) and Sun et al. (2012). As such, static clustering methods are applied to successive subsamples of a fixed size. As explained further later, we refer to these methods as "non-adaptive". As discussed in the statistical literature (see Zivot & Wang 2006, Clark & McCracken 2009), such approaches suffer from various empirical problems. In particular, the choice of the largely arbitrary bandwidth or window size constitutes a crucial assumption that has been proven to significantly affect the results of studies using a fixed window width. For instance, dividing long time series characterized by a time dependence structure may artificially generate outliers that, in addition to the limited timespan of the data within the windows, can severely bias the estimations. Overall, estimations performed over rolling windows can thus provide a misleading picture of the system over time, which calls for alternative approaches specifically designed to account for dynamic structures.

[8]In our empirical application, we define K as the number of existing styles, as identified by Morningstar.

[9]Alternative metrics have also been suggested as potential extensions, such as the Mahalanobis distance (Mao & Jain 1996), the L1 distance (Kashima, Hu, Ray & Singh 2008), and the family of Bregman distances (Banerjee, Dhillon, Ghosh, Merugu & Modha 2007), to quote only a few.

1.3.1.2 Adaptive cluster algorithm

In recent years, alternative statistical clustering techniques have been developed in the machine learning literature to address temporal systems; see, for instance, Yang, Harris, Luo, Xiong, Joachimiak, Wu, Dehal, Jacobsen, Yang, Palumbo et al. (2009), Xu et al. (2014) and more recently Matias & Miele (2017). Here, we rely on the AFFECT algorithm specifically designed to address dynamic clusters. Compared with the previous rolling window methodology, AFFECT allows us to control for the rate at which past proximities are forgotten in an adaptive way. In other words, it allows the data to speak by estimating rather than arbitrarily fixing the optimal level of smoothing in each time step. Another advantage of relaxing the requirement of arbitrary short sequences (24 months in our case) is the ability to directly consider entries and exits in the market, which mitigates survivorship bias more effectively than in the non-adaptive setting in which complete information is required. Moreover, AFFECT avoids data overlaps. Indeed, proximities in the adaptive setting are computed each month using only within-month daily returns.

The AFFECT algorithm is iterative and proceeds in several structured steps. First, as an initialization phase, it computes a matrix of proximities in *t=1* to which is further applied a standard static clustering method (*e.g.*, K-means in our case). In our context, as the initial proximity matrix, we use the dot product matrix of funds' daily net return time series normalized each month, which corresponds to a correlation matrix between the returns. Then, the similarity matrices in each time step are obtained by revising past information thanks to a state-space representation featuring a parameter, namely the "adaptive forgetting factor", which is a function of past similarity matrices. As such, this method assumes that the time-varying observed matrix of proximities ($\mathrm{W_t}$) can be modeled as a linear combination of a true (latent) proximity matrix (Ψ_t) plus a zero-mean noise matrix N_t. Xu et al. (2014) propose a smoothing estimation procedure of Ψ_t as an alternative to the standard Kalman filter procedure. More specifically, this consists of a convex combination of the static estimation W_t through static clustering techniques and the past smoothed proximity matrix ($\hat{\Psi}_{t-1}$):

$$\begin{aligned} &W_t = \Psi_t + N_t \\ &\hat{\Psi}_t = \alpha_t \hat{\Psi}_{t-1} + (1-\alpha_t) W_t \\ &\text{with} \quad \hat{\Psi}_0 = W_0 \end{aligned} \tag{1.2}$$

α_t controls the rate at which past proximities are forgotten and is thus referred to as the "forgetting factor". The forgetting factor's identification is performed through a shrinkage estimation thanks to a block-model representation (see for more details Xu et al. 2014). However, the true value of the α parameter is interpreted by the authors as the "oracle forgetting factor", which requires perfect knowledge of the true proximity matrix (Ψ_t) as well as the noise variance ($\mathrm{N_t}$). Xu et al. (2014) propose an estimator based on the sample counterparts of the theoretical moments on which it depends, which here are simply the sample mean and variance of $\mathrm{W_t}$. However, given that both Ψ_t and $\mathrm{N_t}$ are time varying, the authors suggest using the spatial mean and variance. Finally, because the estimation requires full unknown information about the block structure (cluster structure), it is performed adaptively (in practice, up to three iterations are required to make the process converge).

1.3.1.3 Adaptive vs. non-adaptive cluster algorithms

Table 1.4 provides preliminary insights into the constitution of our clusters when applying these adaptive and non-adaptive approaches. We describe the stability of the clusters by computing two measures. Cluster stability depicts the similarity percentage of associated funds in two consecutive periods. The statistic reaches 79% in the adaptive setting compared with 75% in the non-adaptive one. Cluster switch corresponds to the percentage of funds migrating from one cluster to another for at least two consecutive months. The results show that funds switch styles 15.7% of the time under the adaptive approach and 19% under the non-adaptive one. Overall, these figures are in line with those of Sun et al. (2012) (16.6%) for U.S. hedge funds and Brown & Goetzmann (1997) (17.6%) for U.S. mutual funds. Therefore, they are consistent with the idea that while funds typically follow one long-term global strategy (*i.e.*, their style), they can occasionally migrate to another style. We also report information on our largest and smallest styles in terms of the population of funds.

Eventually, for the overlapping year, we cross-check our classification with that proposed by Morningstar (Table 1.5). A similar exercise is conducted by Sun et al. (2012), who cross-check their hedge fund classification with the one retrieved from Lipper TASS. Overall, the classifications are consistent. Considering the best match (*i.e.*, the highest number of funds in each category that match with a specific style cluster), 57% of the funds grouped together in Morningstar are still together under our clustering algorithm. If we take the cumulative proportion of the two largest matches, we reach 76%.

Table 1.4: Cluster statistics

Cluster descriptive statistics		
	Adaptive	**Non-adaptive**
Cluster Stability	78.76%	74.59%
Cluster Switch	15.73%	19.01%
Average maximum cluster population	24.28%	12.68%
Average minimum cluster population	1.87%	.41%
Overall maximum cluster population	39.78%	24.83%
Overall minimum cluster population	.34%	.03%

Note: Table 1.4 reports cluster descriptive statistics both for the adaptive and the non-adaptive settings. In the former, clusters are determined using the AFFECT algorithm on within-month daily returns. In the later, the clusters are formed using the K-means algorithm on rolling windows of 24 months worth of returns as in Sun et al. (2012). Cluster stability measures the average % consistency between each cluster composition in time t and time t+1. Cluster switch refers to the probability of a fund to change style (for at least two consecutive periods) over our sample. Average and overall maximum and minimum cluster population report information concerning the largest and smallest clusters.

Table 1.5: Cluster cross-tabulation between Morningstar categories and adaptive clusters (December 2016)

Morningstar Reported category	ID 1	5	6	7	14	15	17	20	21		Largest Match	Cum. 2^{nd} Match
Cat 50%MSCI Wld/CD NR&50%MSCI Wld/CS NR	3	4	10	1	5	1	7	11	3		24%	47%
Euronext Paris CAC 40 NR EUR	2	74	13	6	0	0	0	0	0		78%	92%
MSCI AC Asia Ex Japan NR USD	4	1	0	1	59	0	0	1	14		74%	91%
MSCI AC Asia Pac Ex JPN NR USD	2	0	0	6	25	0	1	1	15		50%	80%
MSCI AC Asia Pacific NR USD	6	1	0	28	9	1	0	4	7		50%	66%
MSCI EM Europe NR EUR	8	0	0	0	0	26	0	0	7		63%	83%
MSCI EM NR USD	17	0	0	0	151	0	0	7	26		75%	88%
MSCI EMU NR EUR	25	250	34	4	0	4	0	1	17		75%	85%
MSCI Europe Growth NR EUR	3	1	48	0	0	2	0	0	5		81%	90%
MSCI Europe NR EUR	42	191	161	1	1	24	1	2	43		41%	76%
MSCI Europe Small Cap NR EUR	8	9	21	1	0	18	1	0	13		30%	55%
MSCI Europe Value NR EUR	16	48	4	0	0	3	0	0	6		62%	83%
MSCI Italy NR EUR	0	42	0	0	0	0	0	0	0		100%	100%
MSCI Spain NR EUR	2	57	7	0	0	0	0	0	0		86%	97%
MSCI World Growth NR USD	8	0	1	0	3	10	29	30	10		33%	65%
MSCI World NR USD	84	23	18	0	4	62	209	79	48		40%	56%
MSCI World Value NR USD	16	8	0	0	2	10	32	7	8		39%	58%
MSCI World/Health Care NR USD	7	0	2	0	0	7	18	10	5		37%	57%
MSCI World/Information Tech NR USD	5	0	0	0	0	6	3	22	7		51%	67%
Russell 1000 Growth TR USD	2	0	0	0	0	7	16	20	6		39%	71%
Russell 1000 TR USD	15	1	0	0	0	11	80	21	8		59%	74%
Topix TR JPY	3	3	0	63	0	7	0	3	6		74%	82%
										Mean	**57%**	**76%**

Note: Table 1.5 provides the cross-tabulation of Morningstar reported categories and our clusters following AFFECT (ID) as of December 2016. Largest match reports the % of funds of each Morningstar's category which belong to their closest AFFECT cluster (maximum matching). Cumulative second match reports the percentage of each Benchmark's associated fund belonging to their two closest associated AFFECT cluster.

Equipped with these two procedures (adaptive and non-adaptive clustering), we can compute SDI measures in the cluster of reference for each fund (*i.e.*, the group of peers following the same style) over time. Table 1.6 features the evolution over time (yearly basis) of the computed SDI (monthly average over all funds) based on the adaptive K-means generated from the AFFECT algorithm (column I) and from the non-adaptive K-means (column II). As explained previously, the adaptive setting does not rely on rolling windows and incorporates directly available data on a monthly basis (reducing the loss of data compared with the rolling window approach), thereby allowing us to better distinguish the dynamics of the SDI over time. Yet, common patterns can be observed for both of them. There is a significant increase in commonalities after the explosion of the dotcom bubble (2001–2003), followed by a sharp decrease starting in 2004 and peaking in 2005. A second, more pronounced, rise in commonalities occurs around the financial crisis (2006–2008) and is followed (in the adaptive case) by a period of relatively high distinctiveness around 2009–2010. In the non-adaptive case, the crisis impact is more persistent considering the level of smoothing imposed (24 months). In 2012–2013, the market exhibits higher average distinctiveness, immediately followed by a short yet intense commonality rise during the Chinese financial crisis in late 2015. This table illustrates one caveat of using a fixed window width, which tends to artificially overstate critical periods (dotcom bubble, financial crisis) and understate smaller transitive periods (2009–2010).

1.3.2 Regression analyses

We then use a panel regression analysis setting to shed light on the impact of the SDI on mutual funds' performance and identify its determinants. For the performance analysis, we use both a benchmark model in which the SDI enters linearly and alternative non-linear models to explore more complex relationships. In addition, two performance measures are computed. In the spirit of Sun et al. (2012), we include our distinctiveness variable as a regressor to explain one-step-ahead excess returns as measured by the traditional four-factor alpha. Alternatively, and to ensure consistency with the principle of style-based competition—the notion that funds mainly compete with funds following the same style—we depart from Sun et al. (2012) and construct an auxiliary dependent variable. This variable depicts the relative performance of the fund, namely its performance compared with that of its direct competitors. Because its computation is less straightforward than that for standard excess returns, we follow the testing procedure developed by Ardia & Boudt (2018), which allows for a formal assessment of whether pairs of estimated alphas are statistically different from one another and eventually for each fund, to compute the proportion of peers outperformed. This measure is an adaptation of the well-known "false discovery rate" measure used by

Table 1.6: SDI Dynamics

Year	**Adaptive**	**Non-Adaptive**
1999	0,1273	NA
2000	0,1333	NA
2001	0,1182	0,0721
2002	0,1110	0,0566
2003	0,1142	0,0429
2004	0,1364	0,0465
2005	0,1497	0,0658
2006	0,1287	0,0555
2007	0,1013	0,0588
2008	0,0973	0,0444
2009	0,1109	0,0420
2010	0,1110	0,0444
2011	0,1038	0,0558
2012	0,1285	0,0542
2013	0,1347	0,0615
2014	0,1090	0,0816
2015	0,0920	0,0606
2016	0,1105	0,0376

Note: Table 1.6 reports the average funds' SDI from 1999 to 2016. Column 2 reports the results in the adaptive setting while Column 3 reports those in the non-adaptive setting.

Barras et al. (2010) to analyze peers. Hence, our second performance measure provides the proportion of competitors each fund outperforms. We further explore the link between the SDI and financial performance by looking at the conditional effects. To this end, we include interaction terms and test for a threshold effect by adding a quadratic term. Eventually, in line with the notion of migration risk, we examine how costly the transition toward an innovative strategy could be by including the variation of the SDI in the model.
In a last step, we further extend the analysis to explore the drivers of distinctiveness among EEMFs. We do so by regressing our strategy distinctiveness measure on a set of individual characteristics as well as global factors. At the micro level, this final step describes the various contexts that lead funds to diverge from "the crowd" (Vozlyublennaia & Wu 2017). At the macro level, our study relates to the literature on commonality in the asset management industry (see Bussière, Hoerova & Klaus 2015; Béreau, Casteleyn, Gnabo & Zwinkels 2015) and sheds light on the determinants of the fragmentation and integration phases in the EEMF industry. To that end, we regress our measure of distinctiveness on state-of-the-art fund-specific determinants (see Sun et al. 2012) and cluster- and system wide-level drivers. Table 1.7 defines our main variables. In Section 4.1, we discuss the specification of each model further.

1.4 Results

In this section, we summarize the main characteristics of the EEMF market. Appendix A describes the industry more in depth. The European market displays several traits that provoke the interest of academic researchers. We recall them in the main lines (Table 1.8) before presenting our findings. For instance, mean size (TNA), age, and the four-factor alpha are €234.4 million, 9.5 years, and -0.81% (annually), respectively. These figures are close to those reported by Ferreira et al. (2013) for EU countries[10].

At a structural level, the EEMF market displays a unique combination of integration and fragmentation features compared with the United States. Hence, the European market is integrated, as countries typically apply the laws set by the European Parliament. Yet, the interpretation and implementation of those laws can diverge across countries. Further, European economies can experience different macroeconomic developments or levels of competition, thereby offering contrasting environments to local mutual funds. Regarding fiscal

[10]Ferreira et al. (2013)'s estimates are $251.9 million, 10.6 years, and -0.39% (quarterly).

Table 1.7: Control variables definitions

Fund specific variables	
Fundsize	Natural logarithm of a fund's TNA in bn
Age	Natural logarithm of a fund's age in years
Flow	Change in TNA not explained by the fund's performance
Volatility*	Std of within-month daily net returns
Cluster specific variables	
Cluster-mean*	Average returns of the funds belonging in each cluster
Cluster-size	% of funds belonging in each cluster
Cluster-SDI*	SDI of cluster with respect to the overall population of funds
Cluster-meancap	Average size (in TNA) of funds belonging to the cluster

Note: Table 1.7 reports the definitions of our main fund and cluster characteristics.* indicates that the way to compute the associated variables differ in the non-adaptive setting using 24 months of past returns instead of within-month daily returns.

aspects, for instance, competitive tax laws in Ireland and Luxembourg foster the domiciliation of funds (PWC 2018). At the fund level, some European features are also worth mentioning. Typically, in most European countries, more than 50% of domestic EEMFs' TNA are owned by commercial banks compared with less than 20% in the United States (Ferreira, Matos & Pires 2018) because of differences in legislation between both markets (*i.e.*, the Glass–Steagall Act of 1999 separating banking and asset management activities until 1999 in the United States). Against this background, we aim in this analysis to empirically describe this segment of the mutual funds market and unveil of some of its main features.

Turning to our results, at the macro level, our large database enables us to show that the EEMF population experienced a strong expansion phase from 1999 to 2009 and then stabilized until 2017. Cross-country comparisons exhibit marked heterogeneity in some periods. On the one hand, the Italy and Spain data emphasize limited increases in the equity mutual fund population during the first decade of the sample and a significant decline during and after the European debt crisis. On the other hand, Luxembourg, France, and Germany (to

Table 1.8: European mutual funds descriptive statistics

Fund descriptive statistics			
	Mean	**Median**	**Standard deviation**
Fundsize (€M)	234.48	66.91	663.46
Age (Years)	9.51	8.17	7.52
Flow (€M)	-0.46	-0.03	34.81
Volatility	1.08	0.93	0.65
Alpha 4 Factor (Annualised %)*	-0.81	0.09	0.22
Outperformance (%)*	9.5	0	21.52
SDI*	0.12	0.08	0.12

Note: Table 1.8 reports descriptive statistics concerning our main fund characteristics and performance/distinctiveness measures using the adaptive setting.* Indicates that the way to compute the associated variables differ in the non-adaptive setting using 24 months of past returns instead of within-month daily returns.

some extent) display strong expansion phases until 2009 and resist well during the crisis, whereas the Ireland data show growth in the country's equity mutual fund industry after 2009. During the two decades, our findings emphasize periods of concentration from 2000 to 2003, from 2006 to 2011, and in 2015, marked by a lower number of styles and larger style-based groups, along with periods of fragmentation from 2003 to 2006, from 2012 to 2014, and in 2016 characterized by relatively few populated clusters. Overall, our analysis suggests the existence of six key periods characterized by three attributes: level of concentration versus fragmentation, persistence in cluster composition, and level of distinctiveness (see Table 1.9).

1.4.1 Performance analysis

Table 1.10 displays the estimates of the baseline equation that models the one-step-ahead performance as a function of its SDI and a set of control variables. If the SDI mirrors innovative and skillful managerial talents, we should expect its estimated coefficient to be significant and positive. As described in Section 3, the set of regressors can be organized into two categories: fund-specific and cluster-specific attributes. Column I on the left reports the main results. The parameter estimates are obtained by applying the modified version of Sun

Table 1.9: Overall Market trend

Periods	Concentration levels	Clusters persistence	Innovation levels (Mean SDI)
1999 to 2001	Low	Below Average	High
2002 to 2004	Moderate	Above average	Low
2005 to 2007	Moderate	Below Average	High
2008 to 2011	Extreme	Average	Low
2012 to 2015	High	Average	High
2015 to 2016	Extreme	Below Average	Low

Note: Table 1.9 summarises the main characteristics of the European equity mutual funds across time. Based on a qualitative analysis of Figures 6 to 9, we report the overall concentration of the market (in terms of TNA, number and size of the clusters). Then using Figure 11, we report the persistence in cluster composition (% of funds remaining in the same cluster). Finally, we display the average level of SDI as reported in Table 1.6

et al. (2012)'s procedure[11]. From columns I to V, the results are retrieved using an adaptive setting, namely AFFECT, and within-month daily returns (minimum 18 observations). In column I, we apply GLM to estimate the coefficients. In column II, we apply traditional OLS as an alternative. In column III, the autoregressive term is dropped from the specification. In the literature (see Sun et al. 2012, Ferreira et al. 2013) the dynamic nature of the process is traditionally ignored despite its strong significance, as shown by the results in column I. Columns IV and V report the results using the more traditional absolute measure of performance with the four-factor alpha[12] (Carhart 1997). In the second part of the table (columns VI and VII), we report the results for the non-adaptive setting. The difference from the first set of results (columns I to V) lies in the setting used to recover the clusters and, in turn, the performance measures and SDI. In column VII, we strictly follow the approach of Sun et al. (2012). In column VI, we depart from the original approach by replacing the four-factor alpha with the outperformance ratio. These additional models enable us to compare the influence of the adaptive and non-adaptive approaches for our data.

[11]Specifically, our econometric approach departs from Sun et al. (2012) in four ways. First, the dependent variable is measured by the level of outperformance, as provided by the testing procedure of Ardia & Boudt (2018). Second, the SDI is computed using style-based clusters in an adaptive setting. For these clusters, we apply the AFFECT algorithm. Third, to take care of the support space of the dependent variable that displays the percentage of funds outperformed within the same cluster, we apply generalized linear model (GLM) estimators. The GLM methodology is more efficient than traditional ordinary least squares (OLS) for explaining bounded variables between 0 and 1 (see Papke & Wooldridge 1996). Fourth, we include an autoregressive term to account for potential inertia.

[12]The global factors are retrieved from Kenneth French's database: http://mbtuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

Table 1.10: (Out)Performance and SDI: linear models

Panel Regression

		Adaptive setting					Non -adaptive	
		I Outperformance$_{t+1}$ GLM	**II** Outperformance$_{t+1}$ OLS	**III** Outperformance$_{t+1}$ GLM	**IV** 4 Factor alpha$_{t+1}$ OLS	**V** 4 Factor alpha$_{t+1}$ OLS	**VI** Outperformance$_{t+1}$ GLM	**VII** 4 Factor alpha$_{t+1}$ OLS
Fund Specific	**Outperformance**	.022***	.022***	/	/	/	.770***	/
		(10.82)	(10.82)				(264.00)	
	Performance (4 Factor alpha)	/	/		.009***		/	.934***
					(3.30)			(1062.50)
	SDI	14.31***	14.31***	14.79***	.017***	.017***	4.75***	.015**
		(27.09)	(27.08)	(27.48)	(4.77)	(4.74)	(5.53)	(2.16)
	Fundsize(ln)	-.238***	-.238***	-.240***	.000	.000	.241***	.001***
		(-7.03)	(-7.03)	(-6.97)	(1.03)	(1.06)	(9.87)	(6.14)
	Age(ln)	-.238***	-.238***	-.242***	-.002***	-.002***	-.357***	-.001***
		(-4.74)	(-4.74)	(-4.73)	(6.57)	(-6.59)	(-7.37)	(-2.78)
	Flow	.524	.524	.618	.000	.000	2.91***	.044***
		(0.59)	(0.59)	(0.70)	(0.01)	(0.05)	(3.06)	(3.60)
	Volatility	1.44***	1.44***	1.44***	-.029***	-.030***	.003	-.001**
		(9.74)	(9.74)	(9.64)	(-17.87)	(-17.74)	(0.08)	(-2.13)
Cluster Specific	**Cluster-mean**	-1.59***	-1.59***	-1.59***	-.015***	-.011**	1.57***	.009***
		(-3.39)	(-3.38)	(-3.39)	(-3.00)	(-2.15)	(17.26)	(7.68)
	Cluster-size	.039***	.039***	.040***	.000***	.000***	-.064***	.000
		(6.87)	(6.86)	(7.04)	(4.33)	(4.38)	(-5.52)	(1.40)
	Cluster-SDI	-2.92***	-2.92***	-2.98**	.008***	.008***	.129	.005
		(-12.17)	(-12.16)	(-12.24)	(4.17)	(4.28)	(0.34)	(1.21)
	Cluster-meancap	-3.44***	-3.44***	-3.53***	-.025***	-.025***	-2.22***	-.026***
		(-7.44)	(-7.44)	(-7.51)	(-7.98)	(-7.95)	(-6.39)	(-7.94)
	Time fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Fund clustering effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	# of observations	334,072	334,072	334,072	334,072	334,072	347,082	347,115
	# of funds	4,284	4,284	4,284	4,284	4,284	4,267	4,267
	Avg Obs per fund	78	78	78	78	78.0	81.3	81.4
	R^2	13.48	13.48	10.10	60.60	60.60	77.77	96.70
	Fstat	231.17***	231.17***	167.44***	2,313***	2,313***	7,195***	50,324***

Note: Table 1.10 reports panel regressions results for EEMF's drivers of outperformance at the monthly frequency as follows: $outperformance_{t+1} = \beta_{0,i} + \beta_{1,i}Fund - variable_{i,t} + \beta_{2,i}Cluster - variable_{i,t} + \epsilon_{i,t}$.Column I reports our main results using the outperformance ratio as dependent variable and GLM. In Column II, we test the same model applying traditional OLS. In Column III, the autoregressive term is dropped. Column IV and V adopt similar specifications than in Columns I and III but differ in the construction of the dependent variable as it uses the 4 factor alpha. Finally, as a robustness check, we test alternative versions of models I and IV based on a non-adaptive setting. Variables are as defined as in Table 1.7. The robust t-statistics reported in parentheses are adjusted for fund clustering effect and time fixed effect. *** 1% ; ** 5% ; * 1%significance

Throughout the results, we report both the value of the coefficients and the robust t-statistics, corrected for fund-level clustering, in parentheses[13]. In the lower part of the table, we include the number of observations, number of funds, average observation per fund, pseudo R^2, and F-statistics to test whether all the regressors are jointly equal to zero. The pseudo R^2 is computed by following Verbeek (2008) as the correlation of the actual and predicted values of the dependent variable.

We start the analysis with the main model. We find that the coefficient attached to the SDI displays strong statistical significance because the p-value is below the 1% level when we control for the time fixed effect and adjust the t-statistics for fund-level clustered errors, along with a large number of fund- and cluster-specific attributes. The sign of the coefficient is positive, indicating that EEMFs following a distinctive strategy on average outperform their close peers. Hence, a more singular strategy within styles tends to reflect management skills for EEMFs[14]. Economically, the outperformance ratio is expected to increase by 14.31 basis points after a one standard deviation increase in the SDI. Among the fund-specific attributes, all but fund flows are found to be statistically significant at the 1% level. The signs of the coefficients are in line with those of previous studies that explain financial excess returns. For instance, the marginal effect of *Fund size* is negative. As discussed by Ferreira et al. (2013), the sign can be explained by the scarcity of the best investment opportunities. As the fund grows, marginal investments must be made in less performing assets, eroding the portfolio's overall performance (see Sun et al. 2012 for hedge funds and Berk & Green 2004, Chen, Hong, Huang & Kubik 2004, Vozlyublennaia & Wu 2017 for mutual funds). Another explanation is the existence of the organizational diseconomies of scale attributable, for instance, to the difficulty in processing soft information, or so-called hierarchy costs (see Stein 2002). Likewise, the sign of *Age* is negative. EEMFs perform worse as they age. Put differently, younger funds tend to outperform their older peers. This outcome can be explained by the necessity for a fund that has been newly created, for instance, to display particularly strong performance to challenge existing institutions and attract sufficient capital to expand. This result is consistent with the findings of Ferreira et al. (2013) for funds outside the United States as well as Otten & Bams (2002) for EEMFs. *Volatility* exhibits a positive sign, as more volatile funds tend to outperform their peers (the variable is not significant in Sun et al. 2012). Eventually, the autoregressive term is found to be statistically significant and positive, exhibiting signs of strong persistence in outperformance, consistent with the idea of skilled managers being able to outperform persistently.

[13] Such a correction has become standard in this strand of the literature (see Sun et al. (2012), Ferreira et al. (2013), Vozlyublennaia & Wu (2017)).

[14] To the best of our knowledge, we are the first study to provide empirical evidence on strategy distinctiveness for European mutual funds.

Turning to the cluster-specific characteristics, we first note their high significance. This set of controls aims to account for cluster-level heterogeneity in fund performance. For instance, applying Ardia & Boudt (2018)'s procedure, we might fail to detect any statistical difference in excess returns among all the funds constituting a cluster. In this extreme case, the outperformance ratio would be zero for all the funds in the cluster. Alternatively, a cluster can exhibit marked differences across funds and be associated with a positive average ratio. Pooled together, the data display cluster-level heterogeneity that remains unexplained by the fund-specific attributes and that we aim to capture by including this additional category of variables. All four variables are significant at the 1% or 5% level. *Cluster size* displays a positive sign, indicating that funds tend to outperform peers in large style-based clusters. Conversely, clusters including large funds—as measured by the average AUM per fund—and well-performing funds (estimated by their average net return) tend to lower the value of outperformance. The same holds true for clusters distinct from the overall market. The more innovative the style, the more difficult it is for funds following this style to outperform their cluster peers.
We now consider the remaining columns of Table 1.10 (columns II to VII). For the adaptive approach (columns II to V), we note the strong and robust effect of the SDI on financial performance, measured by either cluster-based outperformance or the more traditional four-factor alpha. We find strong consistency between the results of our main model in column I and those of the alternative models using the adaptive setting (columns II to V). Neither the application of the OLS method (column II) instead of GLM nor the exclusion of the autoregressive term (columns III and V) affects the significance of the coefficients for the outperformance ratio and four-factor alpha. We now compare the results of the adaptive approach (columns I to V) with those of the non-adaptive approach (columns VI and VII). The SDI is positive and significant under both approaches. Nevertheless, the comparison also displays three notable differences. First, the statistical significance is stronger for the SDI under the adaptive approach. Second, for the other variables, the conclusion on the statistical significance of the coefficients alters in some cases. Hence, for the 10 regressors tested, we draw a different conclusion for three of them (column I vs. column VI). The signs of the coefficients also differ, as three parameters display the opposite signs (column I vs. column VI).

In Table 1.11, we explore the non-linear effects in the relationship between the SDI and fund outperformance. To this end, we extend our main model (column I) and add a new regressor constructed as the interaction between our variable of interest (*i.e.*, the SDI) and another variable driving the non-linearity. Each interaction regressor is included in a separate regression. Overall, four types of effects are assessed. *First*, we test whether more innovative strategies are rewarded differently in crisis periods. For this purpose, we consider three crisis indicators: (i) a dummy variable for the European crisis (ECD) that takes the value of 1 from January 2008 to March 2009 and from July 2011 to March 2013[15] and 0 otherwise (column II), (ii) a dummy variable for the global financial crisis (GCD) that takes the value of 1 from December 2007 to May 2009 and 0 otherwise (column III), and (iii) *Vstocxx* that captures market uncertainty in Europe (column IV). *Second*, we separate within-cluster and between-cluster distinctiveness strategies. Specifically, we test whether the outcome on out-performance is different when innovative strategies are implemented within clusters distant from the rest of the industry (column V). *Third*, we investigate the existence of a change in the relationship of interest as funds become more distinct. To this end, we include the SDI variable with a quadratic term (column VI). A positive coefficient for the SDI variable and a negative sign for the coefficient attached to the quadratic term suggests that the beneficial effect of a distinctive strategy dampens as such differentiation increases. In other words, being distinct improves financial performance; however, the marginal effect vanishes as a fund's strategy departs from its peers. *Fourth*, we include the variation of the SDI (column VII) to assess the potential cost of the transition toward more distinctive strategies that are less known by the market and for which fund managers have less expertise. This effect is captured by including the first difference of the SDI in the previous period, ΔSDI_{t-1}.

The results reach significance for the two models that test the effect of the crisis. Each time, the coefficients of the SDI and interaction term are positive, indicating that the effect of the SDI is exacerbated in periods of stress. Hence, relative to normal times, being able to implement an innovative strategy in difficult times enables earning higher excess returns than close competitors on average. The coefficient attached to *Cluster-SDI* $*$ *SDI* is negative and statistically significant, which suggests that the more distinctive a style, the more difficult it is to take advantage of a within-style distinctive strategy. Moving to the next column, the coefficient attached to the quadratic form of the SDI shows negative statistical significance. Therefore, although the effect in our sample is strongly positive, evidence of non-linearity exists, and the marginal effect dies out as the SDI grows. Our results are strongly significant for the parameter attached to ΔSDI_{t-1}. This shows that a rapid shift toward more innovative strategies is associated with lower financial performance consistent with the principle of migration risk.

[15] See https://cepr.org/content/euro-area-business-cycle-dating-committee.

Table 1.11: Outperformance and SDI: Non-linear models

Panel Regressions

	Outperformance$_{t+1}$	I	II	III	IV	V	VI	VII
	Outperformance	.022***	.022***	.022***	.021***	.021***	.020***	0.21***
		(10.82)	(10.83)	(10.81)	(10.45)	(10.56)	(9.96)	(10.12)
	SDI	14.31***	13.32***	13.85***	9.39***	17.53***	26.70***	17.95***
		(27.09)	(22.89)	(25.04)	(17.44)	(22.14)	(26.07)	(27.93)
	SDIˆ2						-23.00***	
							(-12.88)	
	ΔSDI$_{t-1}$							-8.28***
								(-15.01)
	Fundsize(ln)	-.238***	-.239***	-.238***	-.197***	-.233***	-.223***	-.218***
		(-7.03)	(-7.07)	(-7.05)	(-5.85)	(-6.94)	(-6.94)	(-6.50)
	Age(years)	-.238***	-.238***	-.237***	-.266***	-.236***	-.222***	-210***
		(-4.74)	(-4.75)	(-4.74)	(-5.38)	(-4.74)	(-4.53)	(-4.11)
Fund specific	**Flow**	.524	.531	.530	1.04	.495	.449	.126
		(0.59)	(0.60)	(0.60)	(0.98)	(0.56)	(0.51)	(1.18)
	Volatility	1.44***	1.49***	1.50***	1.30***	1.46***	1.64***	1.475***
		(9.74)	(10.00)	(9.88)	(12.04)	(9.84)	(11.26)	(10.16)
	Cluster-mean	-1.59***	-1.61***	-1.58***	1.57***	-1.57***	-1.69***	-1.464***
		(-3.39)	(-3.44)	(-3.38)	(7.45)	(-3.36)	(-3.60)	(-3.11)
	Cluster-size	.039***	.039***	.040***	.049***	.047***	.052***	.050***
		(6.87)	(6.84)	(7.07)	(9.23)	(8.11)	(9.45)	(8.83)
Style Specific	**Cluster-SDI**	-2.92***	-2.92***	-2.88***	-2.92***	-1.64***	-2.99***	-3.051***
		(-12.17)	(-12.20)	(-11.98)	(-12.64)	(-5.33)	(-12.65)	(-12.61)
	Cluster-meancap	-3.44***	-3.49***	-3.36***	-3.08***	-3.58***	-4.13***	-.361***
		(-7.44)	(-7.55)	(-7.25)	(-7.35)	(-7.74)	(-8.94)	(-7.75)
Global	**Vstocxx**				-.126***			
					(-17.87)			
	ECD * SDI		3.85***					
			(4.33)					
	GCD*SDI			4.07***				
				(3.27)				
Interaction variables	**Vstocxx*SDI**				.283***			
					(14.24)			
	Cluster-SDI * SDI					-8.12***		
						(-5.60)		
	Fund clustering effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Time fixed effect	Yes	Yes	Yes	No	Yes	Yes	Yes
	# of observations	334.072	334.072	334.072	333,968	334.072	334.072	334.072
	# of funds	4,284	4,284	4,284	4,284	4,284	4,284	4,284
	Avg Obs per fund	78	78	78	78	78	78	78
	R^2	13.48	13.51	13.50	8.54	13.54	13.82	13.91
	Fstat	231.17***	231.77***	231.57***	2,353***	232.36***	237.94***	239.73***

Note: Table 1.11 reports panel regressions results where our baseline model is extended to account for non-linear effects. Column I and II report the interaction variables between SDI and an European crisis dummy *ECD* (II) or a global crisis dummy *GCD* (III). Column IV reports the results while using a market volatility variable Vstocxx rather than time fixed effect as an interaction variable. Column V uses cluster-SDI as an interaction variable, while column VI includes a quadratic form for the SDI. Finally, Column VII tests for style drift risk. Fund and cluster variables are identical to Table 1.7. The robust t-statistics reported in parentheses are adjusted for fund clustering effect and time fixed effect. ***1% significance; **5% significance; *10% significance.

1.4.2 Determinants of the SDI

Column I in Table 1.12 displays the results for the reference model in which the coefficients are estimated by GLM. The specification includes an autoregressive term and the SDI is recovered using the AFFECT algorithm. Columns II to III are added to test the robustness of our conclusions. Column II, the specification, is the same as in column I, but we drop the autoregressive term. Finally, in column III, the SDI is computed using a rolling window approach as in Sun et al. (2012).

All the fund-level variables, with the exception of net flow, significantly influence the SDI. This result holds in all cases but one at the 1% level and is consistent across the three models. The sign of the coefficient is negative for the variable depicting the size of the fund and its age. These results are consistent with the findings of Sun et al. (2012) on U.S. hedge funds. The interpretations of fund age and size are not straightforward. As it stands, the negative sign of the coefficient indicates that large and old funds are less prone to follow a distinctive strategy. Several mechanisms could be at play to explain this result. For instance, a fund that has developed its own investment strategy identity and is recognized in the market might have greater difficulty innovating. Likewise, when an investment has been made in specific investment skills, the cost of shifting to an alternative strategy might be high. On the contrary, newly established funds have greater latitude to pursue original and innovative ideas. Small funds also face fewer constraints from long-term clients regarding modifying their investment practices. However, these conjectures about the channels through which size and age affect the SDI cannot be formally tested within our framework. Similar to Sun et al. (2012), greater volatility is also found to be less prone to innovation. Turning to the cluster-level variables, *cluster meancap* and *cluster SDI* are shown to be positive contributors, while small clusters seem to be populated by more distinct funds.

Table 1.12: Drivers of SDI

Panel Regression

		Adaptive		Non-adaptive
		I	**II**	**III**
	SDI_{t+1}			
Fund specific	**SDI**	.552***	/	.928***
		(59.19)		(147.81)
	Fundsize(ln)	-.002***	-.005***	-.000***
		(-8.24)	(-7.98)	(-5.32)
	Age(ln)	-.003***	-.006***	-.000***
		(-6.83)	(-6.87)	(-7.25)
	Flow	-.003	-.000	.001
		(-1.04)	(-0.09)	(0.94)
	Volatility	-.003*	-.030***	-.002***
		(-1.69)	(-10.51)	(-15.01)
Style Specific	**Cluster-mean**	.003	.003	-.000*
		(1.57)	(1.30)	(-1.85)
	Cluster-size	-.001***	-.002***	-.001***
		(-29.56)	(-31.70)	(-30.26)
	Cluster-SDI	.032***	.083***	-.007***
		(13.95)	(16.37)	(-7.86)
	Cluster-meancap	.035***	.082***	.001**
		(10.75)	(12.89)	(2.27)
	Time fixed effect	Yes	Yes	Yes
	Fund clustering effect	Yes	Yes	Yes
	# of observations	334.073	334.073	347,115
	# of funds	4.285	4.285	4,267
	Avg Obs per fund	78	78	81.3
	$\mathbf{R}^2$	66.49	45.82	90.14
	Fstat	2,957***	1,266***	15,778***

Note: Table 1.12 reports the panel regressions results for EEMFs' SDI drivers at the monthly frequency: $SDI_{t+1} = \beta_{0,i} + \beta_{1,i}Fundvariable_{i,t} + \beta_{2,i}Clustervariable_{i,t} + \epsilon_{i,t}$ Column I and II report results in the adaptive setting for dynamic panel regression and panel regression respectively. Column III reports the results in the non-adaptive setting. Fund and cluster variables are identical to Table 1.7. The robust t-statistics reported in parentheses are adjusted for fund clustering effect and time fixed effect ***1% significance; **5% significance; *10% significance.

1.5 Conclusion

In this analysis, we describe the main characteristics of the EEMF industry over the past two decades. To this end, we create a novel database from Morningstar comprising 4,957 EEMFs. Equipped with these data, we use a regression approach to test whether distinctive strategies—as measured by the distance of a fund's returns from the average return of its style-based cluster—are significantly associated with excess performance. Next, we explore the determinants of strategy distinctiveness. We contribute to the literature in three main ways. First, we take particular care in handling several econometric issues that have been overlooked in previous work on strategy distinctiveness and fund performance. To this end, we apply a modified version of Sun et al. (2012)'s SDI based on the adaptive clustering approach called AFFECT, which was developed by Xu et al. (2014). This method relies on a limited set of assumptions and allows us to retrieve mutual funds' styles endogenously over time, thereby addressing changes in the number of styles, funds' shifts in style, and the entry and exit of funds. In addition, we apply the recent approach of Ardia & Boudt (2018) to formally test the difference in performance across peers. Second, to the best of our knowledge, we are the first study to assess the impact of strategy distinctiveness in the context of EEMFs. Third, we explore the non-linear relationship between financial performance and strategy distinctiveness.

Our results show that distinctiveness among EEMFs decreased sharply before the crisis before increasing and eventually hitting a new low in recent months. We find a strong, robust, and positive impact of strategy distinctiveness on financial performance. Interestingly, our analysis also unveils the existence of non-linear effects. In particular, we find a decreasing marginal effect as funds become more distinct. Hence, although following innovative and distinctive strategies pays off on average, the effect is actually non-linear and exhibits a threshold level over which being too different becomes useless. Eventually, a fund's size and age are found to be significantly negative determinants of an innovative strategy.

Chapter 2

Portfolio Concentration and Financial Performance: Insights From Domestic and Global Equity Mutual Funds

2.1 Introduction

While traditional portfolio theory suggests that asset managers should hold shares of a wide variety of individual stocks to eliminate the portfolio idiosyncratic risk (Markowitz 1952, Jensen 1969, Statman 1987, 2004), a rapid scrutiny of asset management practices and academic studies cast doubts on whether passive diversified strategies are unanimously viewed as beneficial. In recent years, many study based on information advantage theory found evidence that some skilful managers were able to outperform their peers persistently by concentrating their portfolio[1], or in other words, by deviating from a perfectly diversified portfolio. The main intuition behind portfolio concentration is that managers endowed with superior information and/or superior skills will over-weight their "best-ideas", hence the positions on which they have the strongest beliefs. This debate between concentration versus diversification benefits can be traced back to two major schools of thoughts initiated by John Maynard Keynes and Harry Markowitz respectively (Boyle, Garlappi, Uppal & Wang 2012). To paraphrase, the former advocates to invest relatively larger sums into firms with which one has better familiarity, because managers' knowledge and experience are limited

[1] See for instance, Kacperczyk et al. (2005), Brands et al. (2005), Baks, Busse & Green (2007), Cremers & Petajisto (2009), Huij & Derwall (2011), Amihud & Goyenko (2013), Choi et al. (2017), Fulkerson & Riley (2019)

and should be used to the fullest extent (Keynes, Johnson & Moggridge 1983). As stated in Boyle et al. (2012), others academics and professionals have supported this view such as Loeb (2007) which states that diversification is an admission of inexperience aiming at reaching an average, or Warren Buffet with is well known moto “Never invest in a business you cannot understand”. Yet, portfolio concentration is not limited to asset concentration, as some managers may “bet”” on specific risk factors on which their have supposedly better information to reap excess performance.

However, the way to properly define and evaluate portfolio concentration as well as its link to managerial skill are still very much debated. In the literature, portfolio concentration either refers to stock concentration (stock selectivity), risk concentration (risk factor exposures) or a combination of both. For instance, Baks et al. (2007) and Fulkerson & Riley (2019), measure a portfolio’s stock concentration on absolute terms using its underlying weights distribution. Others, such as Brands et al. (2005) and most notably Cremers & Petajisto (2009) —with the “Active Share”—, define portfolio concentration as the extend to which the fund’s portfolio deviates from its referential benchmark. Regarding the risk concentration, Cremers & Petajisto (2009) define it as a deviation in risk exposures from a referential benchmark (active risk). Alternatively, Amihud & Goyenko (2013) propose to use the R-square stemming from a multi-factor model to evaluate a fund systematic risk exposure, while Kacperczyk et al. (2005) evaluate it on absolute term with the “Industry concentration Index”. Similarly, studies by Merton (1987), Levy & Livingston (1995), Van Nieuwerburgh & Veldkamp (2009) define concentration through the over-weighting of domestic stock (commonly known as the “homebias”).

In this paper, we wish to contribute to the debate by proposing a holding-based methodology to evaluate portfolio concentration, both in stocks and risk factor exposures[2]. However, we depart from the literature in the way we evaluate the later. We propose to determine the risk factors —used to measure the risk concentration— endogenously through a principal component analysis (PCA) on their holdings’ returns. Our main contribution lies in two key characteristics of the measure.

First, given that the factors are determined endogenously we argue that our framework should be more flexible than traditional multi-factor models applied to US or international funds. Indeed, while the identification of risk factors may be trivial for the US market[3], the task becomes less straightforward when considering other geographic areas such as Europe, where

[2] In the spirit of Cremers & Petajisto (2009)

[3] The vast majority of studies concerning managers activeness and or portfolio concentration are based on US mutual funds with domestic strategies (see for instance Kacperczyk et al. 2005, Cremers & Petajisto 2009, Amihud & Goyenko 2013, Kacperczyk, Nieuwerburgh & Veldkamp 2014) and as such use well documented economic-based risk factors tailored for the US market such as industry factors or traditional style factors (SMB, HML, MOM)

investment profiles are much more diverse. For instance, some European managers invest massively in their domestic market, other in foreign European countries, while a preponderant part are global funds investing in many parts of the world[4] (EFAMA 2017). Identifying their set of relevant factors becomes then much more demanding. To circumvent this issue, the general consensus is to group the potential factors in three distinct factor models, namely: style, sector, and country[5]. Yet, as pointed-out by Huij & Derwall (2011), it is unclear whether such models capture distinct dimensions of risk. In fact, industry factors may very well capture the same underlying risk as country or style factors. Therefore, we propose to remain agnostic about the potential unobserved risk factors by relying on statistical ones determined endogenously from the data. To do so, we apply a methodology stemming from the risk budgeting literature[6] which relies on a PCA applied to the portfolio holdings' returns. It allows —for each fund— to effectively uncover uncorrelated risk factors[7], and then retrieve their respective contribution to the portfolio variance (Meucci 2010, Meucci, Santangelo & Deguest 2015, Roncalli & Weisang 2016). Thus, we use our measure to proxy the breadth of the underlying strategy (*i.e.*, the number of independent investment decisions)[8] by distinguishing managers specializing their investments and displaying strong exposure to a single source of risk (*Focused*) from those willing to scale their investment strategy by "betting" on multiple risk factors (*Dispersed*).This definition resonates with the one of Grinold & Kahn (2000) which define strategy breadth as the number of investment decisions based on distinct information sets (i.e. the number of independent forecast). We argue that a higher number of independent forecasts will translate into broader exposures to uncorrelated risk factors, thus justifying our method.

Second, both our measures are absolute and as such do not rely on a pre-specified benchmark. Indeed, benchmark-centric measures are by design highly sensitive to the choice of the referential benchmark and may be misleading if the benchmark is not appropriately chosen (Sensoy 2009), moreover they do not accommodate well with multi-benchmark strategy as demonstrated by Amihud & Goyenko (2013). Therefore, solely relying on the managers portfolio's allocations — and associated returns— to highlight their risk exposures (*Focused* vs *Dispersed*), allows to avoid such limitations. The same logic holds true for the stock concentration measure, thus we follow Baks et al. (2007) and Fulkerson & Riley (2019) and define it

[4]The European market is the largest hub for global funds worldwide. Luxembourg alone, accounts for 1 out of 10 global fund worldwide EFAMA (2017)

[5]See for instance, Fama & French (1993, 1998, 2017), Jegadeesh & Titman (1993), Huij & Derwall (2011), Banegas et al. (2013), Tsai & Wu (2015), Choi et al. (2017)

[6]See Roncalli (2013) for a literature review

[7]Although this property comes at the price of non-interpretability of factors, we are interested in the distribution of portfolio risk not the identification of its sources.

[8]See Grinold & Kahn (2000), Huij & Derwall (2011)

directly through the portfolio weights[9]. Indeed, the portfolio weights distribution is a natural proxy for stock concentration, if managers overweight some positions in the portfolio (*Stock picks*), their concentration level will rise. On the other hand, if they allocate similar weights across each portfolio holdings (*Diversified*), it will decrease.

Our empirical analysis is based on a set of 1,746 equity mutual funds domiciled in Europe for the period 2003Q1-2016Q4. We collect the entirety of our data from Morningstar direct, which is a global leader data provider for the mutual fund industry (Del Guercio & Tkac 2008). More specifically, we retrieve at a quarterly frequency the portfolio holdings composition of each fund. We also extract the daily closing price of each of the portfolio holdings along our time horizon. Then, using the funds portfolio weights attached to each unique position at each quarter, we are able to compute the stock concentration using the Herfindahl Hirschman index (HHI henceforth). Similarly, based on the funds holdings daily log-returns, we are able at each quarter[10] to determine the uncorrelated risk factors impacting the portfolio thanks to a PCA analysis. Finally, using an inverse HHI (HHI^{-1})[11] on the percentage of portfolio variance explained by each factors, we determine how many of them truly impact its risk (effective number of factors). We argue that their numbers is a natural proxy for a fund's strategy breadth and therefore use it as the risk concentration measure[12].

In a last exercise, we investigate the relation between the funds concentration measures and their manager's skill. To do do, we cross our two axes of concentration to distinguish different active strategies and test for their performance. We highlight four distinct categories characterised by the combination of either low or high concentration in stocks (*Diversified* vs *Stock picks*) and risk factors (*Dispersed* vs *Focused*), and a fifth category —labelled *Core*— representing funds neither highly concentrated nor diversified in stocks and risk exposures. Our main results highlight that funds able to pick stocks while spreading their risk expositions (*Dispersed stock picks*) are able to outperform their peers. These results are robust to alternative definitions of both concentration measures and held true during the financial crisis.

The rest of the paper is organized as follows. In Section 2, we briefly summarise the literature. In Section 3, we describe our methodology and data sources. Section 4 presents our results. Robustness analysis is performed in Section 5. Section 6 concludes.

[9]This departs from the stock selection proxy of Cremers & Petajisto (2009), the "Active Share" which measure the extent to which a fund deviate from its benchmark.

[10]Using rolling windows of the preceding two years worth of daily returns

[11]The inverse HHI allows to compute the effective number of elements on which it is applied to (in our case, stocks or risk factors)

[12]Please note that in this case lower values of the measure are representative of concentrated strategies

2.2 Literature Review

While the consensus reached in the academic literature over the last decades points to the inability of active mutual funds managers to cover their costs[13], a recent strand of the literature finds evidence that some skilled managers —concentrating their portfolio— are able to outperform their peers. The rational behind this new strand of the literature is that informational advantages[14] may allow some manager to generate return in excess of their costs. For instance, Kacperczyk et al. (2005) demonstrated that US equity funds' managers concentrating their allocations in few industries where they have informational advantages are able to generates positive abnormal net returns for their investors.[15] Baks et al. (2007) highlight that funds willing to takes large bets on a restricted set of positions outperform their competitors and generate excess net returns for their investors. Furthermore, Cremers & Petajisto (2009) which define the portfolio activeness alongside two axes: stock picking (*Active Share*) and factor bets (*Tracking error*)[16], highlight that US equity funds displaying both high stock selectivity and either low or high tracking errors (*i.e.*, *diversified & concentrated stock pickers* respectively) earn positive net performance compared to their benchmark. Gupta-Mukherjee (2013) builds on Cremers & Petajisto (2009) but rather use an active peer-benchmark to compare the funds performance. This synthetic benchmark represents the composite beliefs (position held by at least two funds) of the funds peers. In response to the body of literature previously detailed advocating for risk concentrated strategies, Huij & Derwall (2011) aim to reconcile the two axioms of the fundamental law of active management[17] which states that activeness is driven by managerial skills (stock selectivity) and the breadth of the underlying strategy (i.e the number of independent investment decisions). Thus, the authors propose to focus on the decomposition of the portfolios' active risk sources, rather than only considering its idiosyncratic risk level (*i.e.*, Tracking error). To do so, they use global mutual funds and analyse their loading on sector, country and style factors. Their findings corroborates that concentrated funds (high tracking error) do outperform, yet that funds concentrated in all three risk sources display significantly higher tracking error than those only exposed

[13]See for instance, Jensen (1969), Gruber (1996), Daniel et al. (1997*b*), Carhart (1997), Malkiel (1995), Wermers (2000)

[14]Better knowledge of the home market (Van Nieuwerburgh & Veldkamp 2009), better expertise on specific markets Kacperczyk et al. (2005), etc.

[15]Similarly,Van Nieuwerburgh & Veldkamp (2009) found evidence than funds concentrated in their country of domicile (domestic strategies) were able to outperform internationally diversified ones. This phenomenon, commonly known as the "homebias"can be explained by informational advantages and learning skills

[16]The former refers to the absolute sum of a fund portfolio's holdings deviation from its benchmark, while the later refers to its active risk (tracking error)

[17]See, Grinold (1989), Grinold & Kahn (2000)

to one or two, hence reconciling tracking error with the breadth of the underlying strategy. More recently, Amihud & Goyenko (2013) proposed a novel measure of risk concentration. They use the R-square (R^2) stemming from a multi-factors regressions on a fund's returns to proxy its idiosyncratic risk exposure[18] (1-R^2), and find that lower R^2 funds outperform their benchmark. Finally, Kacperczyk et al. (2014) and Choi et al. (2017) extend the literature, the former by studying the dynamics of stocks selection and market timing for US funds, the latter by extending the scope of previous studies to engulf institutional investors worldwide. Kacperczyk et al. (2014) uses Daniel et al. (1997*b*) "Characteristic Selectivity" (CS) and "Characteristic Timing" (CT)[19] and find that funds able to pick stocks in expansion and time the market in recession are the overall best, even after fees. Choi et al. (2017) shows that funds concentrated in industries and/or international markets may be optimal.

2.3 Empirical Methodology

2.3.1 Data on funds and funds constituents

We extract raw data regarding EMFs from Morningstar direct. Among key advantages of the database are its comprehensive coverage across countries for funds and securities, as well as the availability of information regarding the funds' historical portfolio holdings, hence allowing us to use a single unified database[20]. We restrict our selection to funds domiciled in Europe[21], which includes 18 countries and 2 microstates[22], and extract the portfolio holdings of 2749 EMFs at a quarterly frequency for the 2003Q1-2016Q4 period. Then, we apply successive filters consistent with the literature (see for instance Kacperczyk et al. 2005, Ferreira et al. 2013), by excluding fund of funds, index tracking funds, funds with less than €1 million of asset under management, less than 5 holding positions, less than 2 quarters of holdings

[18]The higher the R^2 the more closely a fund follows the market, thus the less active it is

[19]CS measures the ability to select stocks that outperform those with similar characteristics, CT measures the ability to time their portfolio weights on specific characteristics

[20]Kacperczyk et al. (2005), Cremers & Petajisto (2009), Kacperczyk et al. (2014) for instance have to merge the CDA and CRSP database which can lead to data mismatch (Zhu 2020)

[21]As previously detailed, Europe offers a privilege laboratory to analyse risk concentration of mutual funds, given the vast diversity of strategies represented (domestic, international, global, etc.)

[22]Austria, Belgium, Denmark, Estonia, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Netherlands, Portugal, Spain, Sweden, Switzerland, UK. Guernesey and Liechtenstein

information, and finally less than 2 years worth of consecutive monthly returns. We further exclude funds invested in less than 75% of common stocks and focus on long-only portfolios. Then, at each quarter we match the funds holdings with the information retrieved on Morningstar regarding the stocks, which to be included, must have complete daily returns information over the previous two years ($\sim$ 506 days)[23]. After this matching procedure, we exclude funds with less than 60% of their portfolio represented. We end-up with 1,746 European EMFs[24].

Information on fund attributes are available at various frequencies, for the purpose of the present study and for the horizon 2001-2016, we collect gross and net funds' monthly returns along with the funds' monthly total net assets (TNA)[25]. We also retrieve the funds' complete managerial history, the inception date of their oldest share, their domicile country and area of investment. With these data at hands, we compute the net expense ratio (NER henceforth) of each fund as in Elton, Gruber & Blake (2013) by subtracting the monthly gross returns of our funds from their monthly net returns. We further compute the turnover ratio by taking the percentage of a fund's TNA represented by the total number of assets purchased or sold between two periods, whichever is less (Amihud & Goyenko 2013). Then, we create a dummy variable to distinguish funds that are managed by a team (*Team* =1) from the one run by a single manager (*Team* =0). Turning to the funds' quarterly net flows, we follow the methodology of Ferreira et al. (2013)[26]. Finally, the abnormal performance is computed with a Fama-French (FF) three factor model (Fama & French 1993), where the factors are computed through MSCI indexes, as done in Huij & Derwall (2011).To obtain correct estimates of each fund's performance level, the first step is to create categories for our funds according to their main investment area. Morningstar uses information available on prospectus to select the funds' principal area of investment. The funds in our sample are distributed into 50 distinct investment areas which can be summarised into 5 main regions (Global, Asia, Europe, US, Emerging). Table 2.1 provides an overview of the number of funds attributed to each investment areas and main investment regions. We observe that the majority of funds in our sample are invested either globally or in European countries

[23]When less than 10 consecutive days are missing we linearly interpolate the price used to retrieve the returns

[24]Closest studies in terms of dataset are the ones of Graef et al. (2019) Franck & Kerl (2013). The former consider 1464 European funds with portfolio holdings at a quarterly frequency over the 2001 2017 period, the later collected portfolio holdings for 4315 European funds yet restricted their sample to the 20005-2009 period at a semiannual frequency.

[25]If multiple shares are available we retrieve the aggregated TNA at the fund level

[26]
$$Flow_{i,t} = \frac{TNA_{i,t} - TNA_{i,t-1} * (1 + R_{i,t})}{TNA_{t,t-1}}$$

Table 2.1: Funds' Investment Area

Investment Areas (IA)	Number of funds	Main Regions
Asia Emerging Mkts	3	Emerging
Asia Pacific	23	Asia
Asia Pacific ex Japan	36	Asia
Asia Pacific ex Japan ex Australia	40	Asia
Belgium	3	Europe
Brazil	5	Emerging
China	20	Asia
China (Greater)	10	Asia
Emerging Europe,Middle East and Africa	4	Emerging
Euroland	150	Europe
Europe	382	Europe
Europe (North)	10	Europe
Europe Emerging Mkts	35	Emerging
Europe ex UK	24	Europe
Finland	9	Europe
France	42	Europe
Germany	26	Europe
Global	439	Global
Global Emerging Mkts	120	Emerging
Hong Kong	3	Asia
India	18	Asia
Italy	15	Europe
Japan	73	Asia
Latin America	10	Emerging
Netherlands	4	Europe
North America	5	US
Russia & CIS	12	Europe
Switzerland	4	Europe
United Kingdom	14	Europe
United States of America	181	US

Note: Table 2.1 reports the investment areas and regions of the funds in our sample as well as the number of funds associated to them. For the sake of clarity, we have excluded from the table IA with less than 3 fund attached to them, which include Africa, ASEAN country, Australia, Austria, BRIC, Czech Republic ,Global ex Euroland, Greece, Indonesia, Korea, Middle East, Middle East & Africa, Norway, Portugal, Singapore, Spain, Sweden, Taiwan, Thailand and finally Turkey

(440 and 691 funds respectively). The rest of the sample is split between funds investing in Asia, Emerging markets and finally in the US (178, 271 and 186 funds respectively). As documented, funds investing in Europe are either following a domestic strategy (for instance Italy with 15 funds or Belgium with 3 funds) or investing in some European countries (150 funds invest solely in "Euroland" countries).
We extract from Morningstar the MSCI indexes of each main region, in order to compute their respective factors. To get each regions size premium (*SMB*) we subtract their MSCI Small Cap index from their overall MSCI index[27]. Similarly, the value premia (*HML*) are computed by subtracting the respective MSCI Growth indexes from their MSCI Value indexes[28]. Finally, the Market premia (*RMRF*) are computed in excess from the European risk free rate. We select the 1-month European Interbank Offered Rate (EURIBOR) rate as a proxy of the short term risk free rate following Auer (2016). Thus, we compute the performance as follows:

$$R_{i,j} - RF = \alpha_{i,j} + \beta_1 RMRF_j + \beta_2 SMB_j + \beta_3 HML_j + \epsilon_f \tag{2.1}$$

where $R_{i,j}$ is the monthly net return of fund i investing in region j, RF[29] is the risk free rate and alpha is the fund abnormal performance. $RMRF_j$ represents the regional markets risk premia, SMB_j and HML_j are the regional small cap and value premia respectively.

2.3.2 Concentration measures

2.3.2.1 Portfolio's stock concentration

In order to asses the stock concentration of our mutual funds' portfolio at a quarterly frequency, we use a measure stemming from the industrial organisation literature, used as a market concentration index: the Herfindahl Hirschman index (HHI). In this study, we ap-

[27]For instance the SMB for global funds is the monthly difference between the MSCI World index and the MSCI World Small Cap index

[28]The momentum factor cannot be retrieved through indexes, which motivated the use of the FF three-factor model, as done by closet study to ours Huij & Derwall (2011)

[29]As in Auer (2016) we propose an alternative test using a constant RF representing the average RF of our entire time span. The results are almost identical and are available upon request

ply the HHI on the portfolio positions ($\boldsymbol{\omega}_{HHI}$)[30] in order to evaluate how "asymmetric" the manager's weighting strategy is. As stated previously, managers may decide to overweight their "best ideas" (core positions) which will mechanically increase the $\boldsymbol{\omega}_{HHI}$, or try to reap diversification benefits by spreading homogeneously their investments, thus decreasing the $\boldsymbol{\omega}_{HHI}$.

$$\boldsymbol{\omega}_{HHIi,t} = \sum_{j=1}^{N} \omega_{i,j,t}^2 \tag{2.2}$$

where $\omega_{i,j,t}$ is a weight of the fund's portfolio i allocated to stock j at time t. N is the total number of unique stock positions composing the portfolio.

2.3.2.2 Portfolio's risk concentration

Computing the funds' risk concentration is less straightforward. Indeed, one can argue that economic-based factors such as industry, country or style factors can be used to proxy the funds risk concentration. Yet, as stated previously, two major issues arise. First, the potential number of eligible factors —when dealing with highly heterogeneous funds— may become very large and difficult to handle. Second, economic-based factors while very intuitive to interpret, are correlated with each others. Therefore, studying the contribution to portfolio risk of each factors may lead to under-estimate the true risk concentration if a sub-group actually proxy the same underlying risk source.

Thus, we propose to uncover uncorrelated risk factors directly from the portfolio holdings. As noted by Meucci (2010), the most natural candidates are the factors stemming from a PCA on the funds covariance matrix "$\boldsymbol{\Sigma}$". PCA, is a dimensionality reducing tool allowing to ease the interpretability of large datasets while at the same time minimising the loss of information. The procedure transforms a set of "N" correlated variables (here stocks) into a set of "N" orthogonal variables called eigen vectors or in our case principal portfolios (PPs, henceforth). Each PP is assumed to represent an unobserved risk factor and is associated to an eigen value which represent the portion of risk the PP accounts for. PPs are ranked according to the magnitude of their eigen values, this way, the first one accounts for as much

[30] As stated previously, we only focus on the portfolio stock positions, which may not sum to 100%. First funds with less than 60% of their portfolio represented are excluded, then to ensure comparability, we re-scale the portfolio positions to sum-up to 100%

risk as possible, and each subsequent PPs for as much of the residual risk as possible[31]. Thus, we decompose the portfolio covariance matrix as follows:

$$\mathbf{\Sigma} \equiv \mathbf{E}'\mathbf{\Lambda}\mathbf{E} \tag{2.3}$$

where $\mathbf{E}$ represent the matrix of PPs and their corresponding eigen values (λ_i) are contained in the diagonal matrix $\mathbf{\Lambda}$ in decreasing order. To implement this method we use for each fund at each quarter, a matrix made of the normalised daily log returns from the portfolio underlying positions over the preceding two years ($\sim$ 506 working days). After having identified the uncorrelated sources of risk arising from each manager's selected set of assets, we are able to analyse their relative contribution for the portfolio variance. As showed by Meucci (2010) and Roncalli & Weisang (2016), the portfolio overall return (R_i) can be expressed as a combination of the original weights ($\omega_{i,j}$) and their associated stocks' returns(r_j) : $R_i = \sum_{j=1}^{N} \omega_{i,j} * r_j$ or as a combination of their PPs with weights : $\tilde{\boldsymbol{\omega}} \equiv \mathbf{E}^{-1}\boldsymbol{\omega}$. Then, the portfolio variance can be expressed through the eigen values as follows:

$$\sigma_i^2 = \sum_{j=1}^{N} \tilde{\omega}_{i,n}^2 \lambda_{i,n} \tag{2.4}$$

where σ_i^2 represents the portfolio variance of fund i, $\tilde{\omega}_{i,n}^2$ is the squared weight attached to the n^{th} PP and $\lambda_{i,n}$ its eigen value. Then, to isolate the % of variance explained by each risk factors (p_n), we divide their contribution by the portfolio variance.

$$p_n = \frac{\tilde{\omega}_n^2 \lambda_n}{\sigma^2} \tag{2.5}$$

While the PPs themselves are not straightforward to interpret[32], their p_n are of great interest to us. Given their additive property we can apply a traditional measure of concen-

[31] See Jolliffe & Cadima (2016) for a detailed review of principal component analysis and its applications

[32] Indeed, the endogenous factors are not identifiable, as such it is not possible to link them to traditional economic-based factors. However, they provide useful information regarding the risk exposures of the portfolio given their uncorrelated nature

tration/diversification in order to evaluate the portfolio risk concentration. In this case, we propose to use an inverse HHI, which conversely to the original index is a measure of diversity. The inverse HHI spans from 1 to N and represents the effective number of factor bets ($\mathbf{P}_{HHI_{i,t}^{-1}}$).

$$\mathbf{P}_{HHI_{i,t}^{-1}} = \frac{1}{\sum_{n=1}^{N} {p_n}_{i,t}^2} \tag{2.6}$$

This measure relates to the breadth of the underlying strategy, defined by Grinold & Kahn (2000) as the number of independent investment decisions, and is an alternative of the model devised by Huij & Derwall (2011). However our methodology, contrary to theirs, does not rely on predefined economic-based risk factors and provides a more flexible framework to study global, international or domestic funds together.

To sum-up, Table 2.2 reports the summary statistics of our main variables. The funds in our sample are at the median worth €221 million and are on average 10 years old. The majority of funds are run by a single manager and their average turnover and expense ratio are around 15.5% and 1.6% respectively. The average quarterly net flows is positive (3.73%) with a standard deviation of 26.86%, reflecting a growth in this segment of the industry over our sample period, but also marked heterogeneity. The funds in our sample invest on average into 86 different stocks, yet some invest in more than 700 assets while others in less that 20. Regarding the portfolio stock concentration ($\boldsymbol{\omega}_{HHI}$), the average level is quite low, with a value of 278.35[33], showing signs that managers tend to diversify across their asset allocations. On the contrary, the majority of funds tends to be mostly exposed to 1 unobserved risk factor ($\mathbf{P}_{HHI^{-1}}$) which support the idea that the average manager specialises its investments in a specific market which overall movement explains most of the portfolio variance. Conversely, some managers are able to bet on up to 5 different unobserved factors. Finally, the average net performance of our funds is around -0.29% (quarterly), which corroborates the general literature consensus concerning the average inability of active managers to cover their costs. Our results are in line with Ferreira et al. (2013), which is to date the most extensive research on mutual funds worldwide. Their European funds, are for instance on average worth $267 million are 12 years old and have a quarterly mean performance of -0,38%.

[33] $\boldsymbol{\omega}_{HHI}$ is bound from 1 to 10000

Table 2.2: Descriptive statistics

	Mean	**Median**	**Std**	**Min**	**Max**
#Assets)	85.80	59.00	92.43	19.00	749.00
Age (years)	10.34	8.83	8.02	0.25	45.51
Alpha (FF)(% Quarterly)	-0.29	-0.40	2.06	-6.40	6.28
$\mathbf{P}_{HHI^{-1}}$	1.08	1.05	0.11	1.00	5.81
$\boldsymbol{\omega}_{HHI}$	278.35	245.22	151.86	39.57	942.34
Flow (%)	3.73	-0.96	26.86	-43.44	171.00
NER(% Annual)	1.60	1.59	0.51	0.19	3.35
Team	0.39	0.00	0.49	0.00	1.00
TNA (€M)	443.72	221.05	669.57	5.11	4222.33
Turnover (% Quarterly)	15.54	13.14	10.12	1.57	54.67
Vol (% 24 months)	4.72	4.38	1.82	2.03	10.80

Note: Table 2.2 reports the descriptive statistics of our main variables over the 2003Q1 to 2016Q4 period. More precisely it reports the mean, median, standard deviation minimum and maximum of the funds' number of assets, age expressed in years, the risk adjusted performance (alpha FF) over the quarter, the stock concentration ($\boldsymbol{\omega}_{HHI}$) and effective number of factor ($\mathbf{P}_{HHI^{-1}}$) , the net flows, the annual net expense ratio (NER), a dummy variable (Team) taking the value 1 if a fund is teamed managed 0 otherwise , the funds total net assets (TNA) and finally their quarterly turnover ratio.

2.3.3 Portfolio concentration and fund's performance

In this section, we propose different models to test the impact of concentration measures on performance, first on a stand-alone basis, then in interaction. Finally, we distinguish five different managerial styles —based on portfolio concentration profiles— and asses their relative skill through their performance level.

To do so, first we estimate a panel regression forecasting the funds' abnormal returns $\alpha_{i,t+1}$ based on the following model:

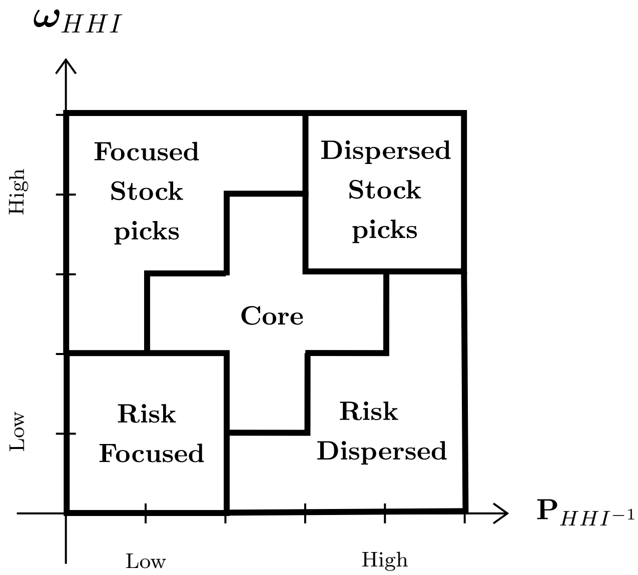
$$\begin{aligned}\alpha_{i,t+1} &= \beta_0 + \beta_{\boldsymbol{\omega}} * \boldsymbol{\omega}_{HHI_{i,t}} + \beta_{\mathbf{P}} * \mathbf{P}_{HHI^{-1}_{i,t}} + \beta_{\boldsymbol{\omega},\mathbf{P}} * (\boldsymbol{\omega}_{HHI_{i,t}} * \mathbf{P}_{HHI^{-1}_{i,t}}) \\ &\quad + \sum_{j=1} \beta_j * Controls_{j,i,t} + \sum_{k=1} FE_k + \epsilon_{i,t+1}\end{aligned} \tag{2.7}$$

where both measures of concentration are first added on a stand-alone basis and then together with their interaction term as in equation (2.7). Our goal is to identify the marginal impact of the different concentration/diversification strategies and their potential interaction. However, both measures could be correlated with other traits of mutual funds, such as their age or their size. Therefore, to isolate our effects, we add a large set of controls. Following the literature (Cremers & Petajisto 2009, Amihud & Goyenko 2013), we include the funds' size computed as the natural logarithm of their TNA (*LN(TNA)*), as well as the funds' age computed as the natural logarithm of the number of years since inception (*LN(Age)*). We further add the funds' net flows (*Flow*), and both turnover ratio (*Turnover)* and net expense ratio (*NER*). Our final controls, are the funds' returns volatility (*Vol*) computed as the standard deviation of their monthly returns over the previous 24 months, and a dummy variable (*Team*) taking the value 1 if the fund is managed by a team and 0 otherwise. Eventually, in line with Choi et al. (2017), we add time, style and investment area fixed effects (FE_k), and as usually done in the literature, we lag all controls by one period and winsorize the continuous variable at the 1% level.[34]

Then, in a second exercise we highlight five different types of active management (managerial styles) and asses their link to subsequent performance in a similar setting than in equation (2.7). The first step, is to define the managerial styles. Similarly to Cremers & Petajisto (2009), we distinguish them along two axes: stock selection ($\boldsymbol{\omega}_{HHI}$) and the strategy breadth ($\mathbf{P}_{HHI^{-1}}$).

[34] To ease the interpretation of our results, we standardized each continuous variable (zero mean and unit standard deviation)

Figure 2.1: Management styles



Note: Figure 2.1 illustrates our managerial grid. The stock concentration ($\boldsymbol{\omega}_{HHI}$), proxy the stock selectivity of the manager by distinguishing funds which overweight some core positions (high) from those spreading homogeneously their investments into multiple assets (low). The number of factors bets ($\mathbf{P}_{HHI^{-1}}$) on the other hand, distinguish funds concentrating on a few factors (low), from those spreading their expositions though multiple risk sources (high). High refers to the 4^{rd} and 5^{th} quintiles of $\boldsymbol{\omega}_{HHI}$ and $\mathbf{P}_{HHI^{-1}}$ at each quarter taking into account all funds available. Similarly, low refers to the 1^{st} and 2^{nd} quintiles of both measures.

Figure 2.1 illustrates five distinct management styles. *Dispersed Stock picks*, refers to funds highly concentrated in stocks, yet able to spread their risk factor exposures, thus effectively broadening their strategy. *Focused Stock picks*, refers to funds concentrated both in stocks and risk factors. *Risk Focused* funds, refers to those following a tight strategy yet homogeneously spreading their stock positions. Conversely, *Risk Dispersed* are funds spreading their risk factor exposures while following an homogeneous stock weighting scheme. Finally, *Core* refers to funds for which neither stocks nor factors concentration characteristics dominate. Equipped with our new categories of active managers (see Figure 2.1), we use dummy variables to uncover their predictive power over their future performance levels:

$$\alpha_{i,t+1} = \beta_0 + \sum_{c=1}^{5} \beta_c * Category_{c,i,t} + \sum_{j=1} \beta_j * Controls_{j,i,t} + \sum_{k=1} FE_k + \epsilon_{i,t+1} \tag{2.8}$$

where $Category_{c,i,t}$ represent the c^{th} dummy variables of management styles associated to each fund at each quarter. To avoid perfect collinearity, we set one parameter (β_c) among the five category dummies equal to zero. The selected parameter provides the base effect against which the four others coefficients should be compared to.

2.4 Results

Before gauging the relationships between our different active management styles and funds' performance, Table 2.3 reports the results of our baseline models. Column I, reports the coefficients of our controls on the funds risk-adjusted performance. *LN(TNA)* is a positive drivers of future performance, in line with Ferreira et al. (2013). Indeed, the authors find that traditional diseconomies of scale recurrent in the US regarding mutual funds are not observed outside of the US, but that in fact there are signs of the opposite. Turning to the positive sign of flows, it supports the findings of Gruber (1996) and its "smart money" effect, which states than investors are able to uncover skilled managers and invest in their funds. Regarding *LN(Age)* and *Turnover*, we highlight that they are negative drivers of performances. The former result, is in line with Ferreira et al. (2013) and support the idea that the performance deteriorates over a fund's lifetime or in other words that young funds tend to outperform older funds. The later, as described in Carhart (1997) tends to highlight the negative impact of excessive trading (excessive transactions cost). Turning to the net expense ratio (*NER*) and the returns volatility (*Vol*), we find that they negatively impact the risk adjusted performance. Finally, we do not find any statistically significant impact of the managerial structure (*Team*) on the funds subsequent performance levels. Columns II, III and IV report respectively the marginal impact of stock concentration, strategy breadth, and their interaction. The positive and highly significant sign of the $\boldsymbol{\omega}_{HHI}$ variable (0.040 at the 1% level) supports the hypothesis than funds which overweight their core positions, are able to reap future abnormal returns (Baks et al. 2007, Fulkerson & Riley 2019). On the other hand, the strongly positive and significant sign of $\mathbf{P}_{HHI^{-1}}$ (0.047 at the 1% level) tends to support the view of the risk budgeting literature (Roncalli 2013, Meucci 2010) which advocates for strategies spreading their risk exposures through multiple risk factors. Finally, when both strategies are examined together, we highlight that the interaction term is highly significant and positive, moreover its coefficient (0.029 at the 1% level) is larger and more significant (1% level) than both coefficients of the stock concentration (0.020 at the 10% level) and strategy breadth (0.025 at the 5% level) measures. Hence, in line with Grinold & Kahn (2000) and Huij & Derwall (2011), we find that skilled strategies are driven together by stocks selection and the breadth of the underlying strategy. Throughout the results, we report both the value of the coefficients and the t-statistics, corrected for fund-level clustering, in parenthesis[35]. In the lower part of the table, we include whether or not we use time, style and country dummy, the number of funds and the number of observations.

Table 2.3: Baseline Models

	Alpha-FF$_{t+1}$	Alpha-FF$_{t+1}$	Alpha-FF$_{t+1}$	Alpha-FF$_{t+1}$
	I	II	III	IV
$\boldsymbol{\omega}_{HHI}$		0.040***		0.020*
		(3.14)		(1.69)
$\mathbf{P}_{HHI^{-1}}$			0.047***	0.025**
			(3.62)	(2.21)
$\boldsymbol{\omega}_{HHI}$*$\mathbf{P}_{HHI^{-1}}$				0.029**
				(2.44)
LN(TNA)	0.051***	0.053***	0.051***	0.052***
	(4.97)	(5.19)	(4.91)	(5.07)
LN(Age)	-0.052***	-0.054***	-0.053***	-0.055***
	(-4.74)	(-4.89)	(-4.80)	(-5.03)
Flow	0.165***	0.165***	0.165***	0.165***
	(20.52)	(20.56)	(20.60)	(20.64)
Turnover	-0.065***	-0.058***	-0.063***	-0.058***
	(-8.13)	(-7.24)	(-7.94)	(-7.29)
NER	-0.028***	-0.030***	-0.030***	-0.030***
	(-3.21)	(-3.38)	(-3.45)	(-3.42)
Team	0.026	0.025	0.027	0.025
	(1.42)	(1.39)	(1.50)	(1.37)
Vol	-0.093***	-0.104***	-0.087***	-0.093***
	(-4.22)	(-4.69)	(-3.96)	(-4.19)
Time fixed effect	Yes	Yes	Yes	Yes
Country fixed effect	Yes	Yes	Yes	Yes
Style fixed effect	Yes	Yes	Yes	Yes
Number of Funds	1746	1746	1746	1746
Number of observations	47934	47934	47934	47934

t-statistics in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Note: Table 2.3 reports the baseline panel regression of the net performance measure (alpha-FF) on lagged control variables. *LN(TNA)* is the natural logarithm of the fund total net assets. *LN(Age)* is the natural logarithm of the fund age expressed in years. *Flow* designate the net fund level flow in % of the lagged value of TNA. *NER* and *Turnover* represent respectively the net expense and turnover ratios. *Team* is a dummy variable taking the value 1 if a fund is team managed, 0 otherwise and Vol represents the funds' returns volatility over the previous 24 months . Finally $\boldsymbol{\omega}_{HHI}$ and $\mathbf{P}_{HHI^{-1}}$ are the measure of stock and risk concentration respectively. We include time, country and style fixed effect, the t-statistics are based on standard errors clustered at the fund level.

[35] Such correction has become standard in this literature (see Cremers & Petajisto 2009, Amihud & Goyenko 2013)

2.4.1 Fund performance: stocks selection vs strategy breadth

So far, we have used concentration measures as continuous variables to asses their impact on performance. To dive deeper into our analysis, in this section we propose to analyse the performance of funds according to their bivariate distribution along $\boldsymbol{\omega}_{HHI}$ and $\mathbf{P}_{HHI^{-1}}$ as represented by the managerial grid in Figure 2.1. Table 2.4 reports for the entire sample, the average performance (*alpha-FF*) of funds according to their concentration measures quintiles. At each quarter, we double sort funds according to both axes and take the time-series average of their performance to compute their overall level over the period 2003Q1-2016Q4. The first results we highlight is that concordant with the literature, on average active funds are unable to covers their costs as demonstrated by the majority of results being significant and negative. Yet, a closer look at funds which overweight their best idea the most while spreading their risk across multiple risk sources, highlights that some skilled managers are able to generate significant positive performances (at the 5 % level) persistently (Kacperczyk et al. 2005, Cremers & Petajisto 2009, Amihud & Goyenko 2013, Kacperczyk et al. 2014). Given the time span of our study (2003Q1-2016Q4) and the number of funds considered we argue that the persistence of the aforementioned category's outperformance makes a strong case against performance being solely due to luck rather than managerial skills.[36] Secondly, for an equivalent high number of factor bets, funds which overweight their core stocks earn an adjusted performance of more than 2% in excess from those spreading their asset allocation homogeneously. Similarly, for an equivalent high level of stock concentration, funds able to bet on multiple factors earn on average 1.6% more than those focusing on 1 factor. These results further corroborates the interaction we observed in Table 2.3 and supports Grinold & Kahn (2000) views on active management. Indeed, stock selection in combination with strategy breadth are key drivers of future abnormal performance.

2.4.2 Active management and performance

Building on the previous section, we will test for the performance of each managerial styles identified in Figure 2.1. To do so, at each quarter we create 5 distinct dummy variables representing each different style (*Focused Stock picks*, *Dispersed Stock picks*, *Risk focused*,

[36]See Barras et al. (2010) for an extensive discussion of managerial luck versus skill

Table 2.4: Alpha-FF EEMFS 2003Q1-2016Q4

$\boldsymbol{\omega}_{HHI}$ quintiles	$\mathbf{P}_{HHI^{-1}}$ quintiles: **Low**	**2**	**3**	**4**	**High**	**High-Low**
Low	-1.785***	-1.5***	-1.571***	-1.625***	-1.601***	0,184
	(-14)	(-11)	(-11)	(-12)	(-6.4)	(0.65)
2	-1.973***	-1.358***	-1.22***	-1.16***	-0.671***	1.302***
	(-12)	(-11)	(-7.3)	(-5.8)	(-2.6)	(4,2)
3	-1.888***	-1.238***	-.360*	-.403**	-0.140	1,748***
	(-11)	(-7.5)	(-1,9)	(-2.1)	(-0.51)	(5.4)
4	-1.645***	-1.305***	-.402**	-0.315*	0.321	1.966***
	(-8.5)	(-5.8)	(-2)	(-1,7)	(1.3)	(6.3)
High	-0.972*	-1.128***	-.941***	.108	0.631**	1.603***
	(-1.9)	(-3.8)	(-2.9)	(0.39)	(2.1)	(2.7)
High-Low	.813	0.382	0.630*	1.733***	2.231***	
	(1.6)	(1.1)	(1.8)	(5.5)	(5.7)	

t-statistics in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Note: Table 2.4 reports the mean annualised alpha-FF (%) of all funds sorted by their $\boldsymbol{\omega}_{HHI}$ and $\mathbf{P}_{HHI^{-1}}$ from 2003Q1 to 2016Q4 (at each quarter).

Risk Dispersed and *Core*)[37]. Table 2.5 reports the results.

In Column I, the category *Dispersed Stock picks* is dropped. In accordance, the coefficients of all four remaining dummies are expressed with respect to it. All four categories, are shown to have a significant negative impact on future performance. More strikingly, all coefficients are significant at the 1% level. The sign is negative, that is, all four categories perform significantly worse than the one representing funds with high stock selectivity and high strategy breadth. For instance, the predicted differential in abnormal returns between our referential category and the *Risk Focused* one (low stock selection and breadth) is about 0.12%. These results suggest that some funds able to keep a highly concentrated portfolio (i.e small portfolio) while keeping a high level of diversity (large strategy breadth) are able to alleviate the negative impact on portfolio diversification imposed by such asset concentration and outperform the rest of the market.

Turning to Column II, we dropped our *Core* category represented by active funds which nei-

[37] More precisely, for instance funds at each quarter who belong simultaneously the 4th or 5th quintiles of stock selectivity and number of effective factor are labelled “Dispersed stock picks” and the associated dummy takes on the value 1. The rest of the categories are highlighted in Figure 2.1

Table 2.5: Active managerial categories and performance

	Alpha-FF$_{t+1}$ I	Alpha-FF$_{t+1}$ II	Alpha-FF$_{t+1}$ III	Alpha-FF$_{t+1}$ IV	Alpha-FF$_{t+1}$ V
Dispersed Stock Picks		0.081***	0.108***	0.145***	0.118***
		(3.01)	(3.62)	(5.01)	(4.03)
Core	-0.081***		0.027	0.064***	0.037*
	(-3.01)		(1.16)	(2.76)	(1.67)
Focused Stock Picks	-0.108***	-0.027		0.037	0.010
	(-3.62)	(-1.16)		(1.38)	(0.41)
Risk Disperse	-0.145***	-0.064***	-0.037		-0.027
	(-5.01)	(-2.76)	(-1.38)		(-1.19)
Risk Focused	-0.118***	-0.037*	-0.010	0.027	
	(-4.03)	(-1.67)	(-0.41)	(1.19)	
Controls	Yes	Yes	Yes	Yes	Yes
Time fixed effect	Yes	Yes	Yes	Yes	Yes
Country fixed effect	Yes	Yes	Yes	Yes	Yes
Style fixed effect	Yes	Yes	Yes	Yes	Yes
Number of funds	1746	1746	1746	1746	1746
Number of observations	47934	47934	47934	47934	47934

t-statistics in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Note: Table 2.5 reports the panel regressions of the funds' performance on controls variables (defined as in table 2.3) and managerial category dummies defined as in Figure 2.1. We include time, country and style fixed effect and t-statistics are based on standard errors clustered at the fund level .

ther excessively concentrate nor diversify their stock or risk exposures. We see that results a more mixed, *Core* are significantly out-performing *Risk Dispersed* and *Risk Focused strategies* at the 1 and 10% respectively. Yet, as stated previously they under-perform *Dispersed stock picks* and are not significantly different from *Focused stock picks*. The later category is dropped in Column III, we do not find any significant differences with respect to the rest of its peers at the exception of *Dispersed Stock picks*, which it under-performs by a differential of approximately 0.11% at the 1% level. Finally, Column IV and V respectively reports the results when the *Risk Dispersed* and *Risk Focused* category are dropped. Overall, these results show that active mutual funds do implement strategies distinct from one another. Moreover, we uncover evidence than those able to concentrate their allocation on their best ideas (core stocks) while loading on multiples risk sources are the most skilled and outperform significantly their peers. This finding once again supports Grinold & Kahn (2000) laws of active management.

2.5 Extensions and Robustness

In this section, we conduct several robustness tests on the benchmark model with alternative measures of both our stock selection and strategy breadth. Then, in a final setting, we test for the performance of our active managerial categories during the financial crisis.
First, we propose an alternative definition of or stock concentration measure. We implement the normalized HHI (nHHI) on the portfolio's' weights as follow:

$$\boldsymbol{\omega}_{nHHI_{i,t}} = \frac{\sum_{j=1}^{N} \omega_{i,j,t}^2 - \frac{1}{N}}{1 - \frac{1}{N}} \tag{2.9}$$

where N is fund's i total number of positions j at time t. The nHHI ranges from 0 to 1 and has the particularity to be invariant to the total number of assets. As such, it only asses the weights distributions. If a fund equally weights each position in its portfolio, its nHHI will be 0, conversely the less homogeneous the weighting scheme is, the closer to 1 the nHHI will be. Table 2.6 reports the results and confirms the main findings of Table 2.5. *Dispersed Stock picks* are still outperforming all others categories (albeit at the 5% level with respect to the *Focused Stock picks*). However, we no longer observe any statistical differences between the rest of the categories. Therefore, we conclude that the actual number of assets composing the portfolio has an impact on its subsequent performance levels.[38]
Second, we propose an alternative measure of the strategy breadth. We use the "Effective number of bets" measure devised by Meucci (2010). It closely relates to an inverse HHI in the sense that it measures the actual number of factors a fund is exposed to. However, it is based on the Shannon entropy, a measure used in information theory to quantify the "uncertainty" linked to a random variable. The measure is defined as follows:

[38] When using the number of assets directly as a concentration measure, similar results arise as the *Dispersed Stock picks* are once again outperforming all others categories. Yet, *Risk Focused* are no longer significantly under-performing the *Core* category. The table is available upon request

Table 2.6: Alternative $\boldsymbol{\omega}_{HHI}$: Normalized HHI

	Alpha-FF$_{t+1}$ I	Alpha-FF$_{t+1}$ II	Alpha-FF$_{t+1}$ III	Alpha-FF$_{t+1}$ IV	Alpha-FF$_{t+1}$ V
Dispersed Stock Picks		0.077***	0.076**	0.081***	0.079***
		(3.03)	(2.45)	(2.67)	(2.88)
Core	-0.077***		-0.001	0.004	0.002
	(-3.03)		(-0.04)	(0.18)	(0.11)
Focused Stock Picks	-0.076**	0.001		0.005	0.003
	(-2.45)	(0.04)		(0.17)	(0.13)
Risk Disperse	-0.081***	-0.004	-0.005		-0.002
	(-2.67)	(-0.18)	(-0.17)		(-0.09)
Risk Focused	-0.079***	-0.002	-0.003	0.002	
	(-2.88)	(-0.11)	(-0.13)	(0.09)	
Controls	Yes	Yes	Yes	Yes	Yes
Time fixed effect	Yes	Yes	Yes	Yes	Yes
Country fixed effect	Yes	Yes	Yes	Yes	Yes
Style fixed effect	Yes	Yes	Yes	Yes	Yes
Number of funds	1746	1746	1746	1746	1746
Number of observations	47934	47934	47934	47934	47934

t-statistics in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Note: Table 2.6 reports reports an alternative setting for Table 2.5 where the measure of stock selectivity in based on a normalised HHI on stock weights. We include time, country and style fixed effect and t-statistics are based on standard errors clustered at the fund level.

$$ENB = exp\left(-\sum_{j=1}^{N} p_n ln p_n \right) \tag{2.10}$$

where p_n are the proportion of portfolio variance explained by the exposure of the fund to its n^{th} factor.

Table 2.7 confirms our main findings (see Table 2.5). Finally, as an extension, we propose to analyse the predictive power of our managerial categories on performance during the crisis. To do so, we set-up a cross-sectional analysis of the funds' performance level during the crisis on their managerial style taken prior to the crisis. First, controls ($Controls^{pre\text{-}cs}$) are taken

Table 2.7: Alternative $\mathbf{P}_{HHI^{-1}}$: Effective number of bets

	Alpha-FF$_{t+1}$ I	Alpha-FF$_{t+1}$ II	Alpha-FF$_{t+1}$ III	Alpha-FF$_{t+1}$ IV	Alpha-FF$_{t+1}$ V
Dispersed Stock Picks		0.090***	0.111***	0.143***	0.120***
		(3.28)	(3.78)	(4.89)	(4.13)
Core	-0.090***		0.021	0.052**	0.030
	(-3.28)		(0.88)	(2.16)	(1.36)
Focused Stock Picks	-0.111***	-0.021		0.032	0.009
	(-3.78)	(-0.88)		(1.18)	(0.39)
Risk Disperse	-0.143***	-0.052**	-0.032		-0.022
	(-4.89)	(-2.16)	(-1.18)		(-0.98)
Risk Focused	-0.120***	-0.030	-0.009	0.022	
	(-4.13)	(-1.36)	(-0.39)	(0.98)	
Controls	Yes	Yes	Yes	Yes	Yes
Time fixed effect	Yes	Yes	Yes	Yes	Yes
Country fixed effect	Yes	Yes	Yes	Yes	Yes
Style fixed effect	Yes	Yes	Yes	Yes	Yes
Number of funds	1746	1746	1746	1746	1746
Number of observations	47934	47934	47934	47934	47934

t-statistics in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Note: Table 2.7 reports reports an alternative setting for Table 2.5 where the measure of strategy breadth is based on the effective number of bets from Meucci (2010).We include time, country and style fixed effect and t-statistics are based on standard errors clustered at the fund level .

as an average of the three quarters preceding the financial crisis (2007Q1-2007Q3). Second, funds are deemed to belong to a specific managerial category ($Category^{pre\text{-}cs}$) if they were associated to it for more than 50% of their pre-crisis observations. Finally third, we compute the cumulative abnormal performance of each fund over the 2007Q4-2008Q4 ($alpha\text{-}FF_{cs}$) to perform our analysis as follows:

$$alpha\text{-}FF_{cs,i} = \sum_{c=1}^{5} \beta_c * Category_{c,i}^{pre\text{-}cs} + \sum_{j=1} \beta_j * Controls_{j,i}^{pre\text{-}cs} + \sum_{k=1} FE_k + \epsilon_i \qquad (2.11)$$

Table 2.8 reports the results. The category *Dispersed Stock picks*, is as previously observed more predictive of future out-performance than the rest of the market. Indeed, the differential in predicted performance with respect to the *Risk Dispersed* and *Risk Focused* is 0.18% and 0.24% respectively, almost twice the magnitude than the ones previously observed over the entire sample. However the results are no-longer statistically significant with respect to *Focused Stock picks* and only significant at the 10% level compared to *Core* active funds. Finally, we conclude that funds which were diversified in their factor exposures while concentrated in their stock selection prior to the financial crisis, fared better (or at least as well) than their peers after its outbreak.

Table 2.8: Active managerial categories and performance: financial crisis.

	Alpha-FF$_{cs}$	Alpha-FF$_{cs}$	Alpha-FF$_{cs}$	Alpha-FF$_{cs}$	Alpha-FF$_{cs}$
	I	II	III	IV	V
Dispersed Stock Picks		0.126*	0.094	0.179**	0.240***
		(1.77)	(1.25)	(2.39)	(3.34)
Core	-0.126*		-0.032	0.053	0.114*
	(-1.77)		(-0.43)	(0.72)	(1.66)
Focused Stock Picks	-0.094	0.032		0.085	0.146**
	(-1.25)	(0.43)		(1.05)	(1.97)
Risk Disperse	-0.179**	-0.053	-0.085		0.061
	(-2.39)	(-0.72)	(-1.05)		(0.86)
Risk Focused	-0.240***	-0.114*	-0.146**	-0.061	
	(-3.34)	(-1.66)	(-1.97)	(-0.86)	
Controls	Yes	Yes	Yes	Yes	Yes
Country fixed effect	Yes	Yes	Yes	Yes	Yes
Style fixed effect	Yes	Yes	Yes	Yes	Yes
Number of observations	654	654	654	654	654

t-statistics in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Note: Table 2.8 reports an alternative setting for Table 2.5 with a cross-sectional analysis of the funds performance during the crisis. Controls are taken as an average of the three quarters preceding the financial crisis (2007Q1-2007Q3). Funds are deemed to belong to a specific managerial category if they were associated to it for more than 50% of their pre-crisis observations. Finally third, $alphaFF_{cs}$ is cumulative abnormal performance over the 2007Q4-2008Q4 period. We include country and style fixed effect

2.6 Conclusion

We propose an empirical strategy —based on portfolio holdings data— to analyse the portfolio concentration of global and domestic European funds. Against this background, the question we ask is whether funds concentrating their holdings and/or their risk factor exposures are skilled and outperform their peers persistently.

To address this question, we rely on recent developments from the risk budgeting literature to asses the funds' strategy breadth, using a principal component analysis on their holdings returns. Overall, our procedure enables to distinguish —for domestic and global funds alike— the actual number of factor bets made by each fund, without making any assumptions regarding the factors themselves. Equipped with a dataset of 1,746 European equity mutual funds from 2003Q1 to 2016Q4, we highlight five managerial categories based on portfolio concentration characteristics. Namely, funds focusing their strategy (low number of factor bets) while either spreading their asset allocations or concentrating it (*Risk Focused* and *Focused Stock picks*, respectively), funds broadening their strategy by betting on multiple factors and either homogeneously or heterogeneously spreading their allocation between assets (*Risk Dispersed* and *Dispersed Stock picks*, respectively) and finally funds neither excessively concentrating nor diversifying their stock and factor exposures (*Core*).

Our results show that both measures (stock concentration and strategy breadth) are significant positive drivers of risk-adjusted performance when we control for state-of-the-art characteristics, yet that their interaction has an even greater predictive power. If we break down the set of mutual funds into five subgroups with respect to their stock and risk concentration profiles, we show that funds concentrating their stock allocation while spreading their risk exposures are outperforming their peers. This finding is robust to a battery of robustness checks in which we alter the definition of our main measures. These results highlight that funds able to spread their risk exposures may increase their portfolio concentration without worsening portfolio diversification to a point were it would become detrimental. Finally, we show evidence that our results still holds in period of financial turmoil by studying the financial crisis.

Chapter 3

SRI Mutual Funds' Performance and Investment Universe

"Let's say you have two equally competent portfolio managers, one of whom can select investments from an unconstrained universe, but the other is restricted to a subset of that universe. Our intuition says that the former will outperform the latter" R. Huebscher (Advisor perspectives), *The Disappointing Reality about ESG Fund Performance* (2019)

"Andrew Rudd's inescapable conclusion that the integration of environment, social or governance (ESG) criteria in investment processes must worsen portfolio diversification appears to be academic wisdom since nearly thirty years, but is it right?" A.G.F. Hoepner, *Portfolio Diversification and Environmental, Social or Governance Criteria: Must Responsible Investments Really Be Poorly Diversified?* (2010)

3.1 Introduction

The conventional "wisdom" reflected in the above quotes that investment performance should be adversely affected by imposing (ethical) criteria on portfolio composition has been at the heart of the debate in the academic world and the financial industry on socially responsible investments (SRIs)[1]. The underlying idea behind such a view is that a restricted universe

[1]The Forum for Sustainable and Responsible Investment defines SRI as "an investment discipline that considers environmental, social and corporate governance (ESG) criteria to generate long-term competitive financial returns and positive societal impact"(https://www.ussif.org/sribasics). Therefore SRI funds may include — yet are not restricted to — thematic funds and impact funds. The former invest in equity belonging to a common theme (as such are usually broader than sectoral funds), the later are funds who solely invest

might worsen portfolio diversification along with portfolio efficiency (*e.g.*, Markowitz 1952, Barnett & Salomon 2006, Renneboog et al. 2008*b*).[2] In this study, we challenge this narrative, which, if false, could provide a misleading picture of this segment of the market and lead to pervasive consequences, such as making end-user investors underestimate what they should require from their manager compared with the regular strategy, or leading them to overestimate the sacrifice they should accept in terms of risk-adjusted performance in order to meet their ethical goals.[3]

Our main claim is that most fund managers operate, in practice, a restricted list of eligible assets embedded in their reference benchmark(s). Such a list defines their style, which provides them with a simple tool to communicate their general strategy to their clients, eases the performance comparison with relevant peers, and enables them to concentrate their attention on a set of assets they know well (Brown & Goetzmann 1997, Almazan, Brown, Carlson & Chapman 2004, Barberis, Shleifer & Wurgler 2005).[4] From this perspective, the question we ask is whether restrictions in an investment universe imposed by sustainability considerations can be compensated for in other areas. For instance, a manager of an SRI multi-strategy mutual fund investing in large cap, mid-cap, and small cap alike could eventually have more stock opportunities than a manager of a purely "unrestricted" (*i.e.*, without SRI restrictions) small-cap strategy.

This study extends the existing literature on SRI firms (see, among others, Bauer, Koedijk & Otten 2005, Bauer, Derwall & Otten 2007, Gregory & Whittaker 2007, Renneboog et al. 2008*a*,*b*, El Ghoul & Karoui 2017). Since decades, the limit of the available investment

in equity linked to companies whose main goal is to generate a quantifiable beneficial impact on social or environmental issues. However, given the nature of impact funds, their vast majority are not publicly listed

[2]According to Markowitz's seminal work on modern portfolio theory in 1952, a restricted universe such as an ESG-screened universe, cannot be more diversified than a conventional universe, because the former is nested in the latter. This reasoning, however relies on the fact that "conventional" universe—in our case the ones of alternative investments strategies that do not include the SRI nature of assets—are unrestricted, which is not the case in practice.

[3]Massive inflows in the SRI industry along with marked changes in investors' expectations have drawn increasing attention toward the financial performance of this segment of the market over the years. As noted in Derwall, Koedijk & Ter Horst (2011), while investors were initially inclined to accept losses in financial performance in exchange for non-financial utility gains derived from SRI investments, changes have been observed in this industry. SRIs can be seen now as "a 'profit-seeking' approach that accommodates investors in their pursuit of traditional financial goals." Evidence on the importance of financial performance in SRIs can also be found from the early nineties in the survey of Rosen, Sandler & Shani (1991). In this context, assessing what a good manager who properly accounts for such constraints should be able to deliver to his/her clients is of utmost interest for market participants.

[4]As discussed in Almazan et al. (2004), part of the restrictions that face an investment fund—labeled as non-fundamental restrictions—do not depend on shareholder approval; they capture the set of constraints that investors and managers consider as necessary to best define the fund's investment style.

universe has been pivotal in the discussion on SRI screening. However, sound empirical evidence on the matter is missing. One reason for this lacuna could be technical: while several features of investment funds can be directly observed, such as their concentration or portfolio size (*i.e.*, shareholding), the universe of eligible assets does not share this property—at least not fully. While not insurmountable, quantifying the investment universe for all funds needs to be treated with care and requires a few assumptions. We propose to address this issue in this study by introducing an original methodology to uncover each funds' investment universe size. To do so, we exploit granular information about benchmark composition along with portfolio holdings of funds. The underlying idea is to associate each fund to one or several benchmarks depending on their actual portfolio holdings. The set of associated benchmarks defines their style (*e.g.*, US growth and Large Cap Financial) and constitutes the first step of our methodology. Once *de facto* styles are identified for each fund, we aggregate the entire set of assets included in each associated benchmark to recover the investment universe. For example, a relatively simple case is one of a mutual fund investing 100% of its assets in a "US growth" benchmark. Here, we would consider that its universe could be approximated by the number of stocks included in the US growth benchmark. The picture is more blurred when funds do rely on multiple benchmarks. Simple aggregation is an option, but it could lead to misleading cases. For instance, if a fund holds 99% in a small benchmark and only a 1% in a large benchmark, considering all assets populating the latter as being part of its universe could inflate the actual number of assets this fund is actually considering for investing. Our design deals with this caveat by considering how important a benchmark is for a fund. Such information is subsequently featured in the aggregation process wherein more weight is given to assets attached to the main benchmark; inversely, lower importance is attached to those in the secondary benchmarks. Eventually, we obtain a novel and simple measure of a style-based investment universe, labeled *IU Score*, available for each fund at each time period.

Our empirical analysis builds on a micro-level dataset of 2,039 US equity mutual funds for the period 2012Q4-2018Q4. All data are collected from Morningstar. Morningstar is considered among the most reliable and critical sources of information for the mutual fund industry (Del Guercio & Tkac 2008).[5] The extant literature is rife with anecdotes or evidence of how important its information is to professionals. Ben-David, Li, Rossi & Song (2019), for instance, emphasize the sound recognition of Morningstar's expertise and its reputation as an independent agency, noting that "investors [...] take Morningstar's advice at

[5]The reputation of Morningstar is well illustrated in Del Guercio & Tkac (2008) by the following quote: "[T]he brand that has emerged as dominant in the 1990s is not Fidelity, Putnam or even Merrill Lynch but instead is Morningstar. Pozen & Crane, The Mutual Fund Business (1998), p. 75"

face value." We retrieve three key ingredients for our empirical analysis from Morningstar. First, Morningstar classifies US equities into common benchmark categories based on their size tilt (small-cap, mid-cap, or large-cap), value tilt (value, blend, or growth) (these two categories constitute the so-called Morningstar "style-box"), and industry tilt (Communications, Consumer Defensive, Consumer Cyclical, Energy Limited Partnership, Equity Energy, Equity Precious Metals, Financial, Health, Industrials, Infrastructure, Natural Resources, Real Estate, Technology, and Utilities Miscellaneous Sector).[6] We possess holdings for all these benchmarks at the quarterly frequency. Second, we collect the fund's stock holdings for all our mutual funds at a quarterly frequency. By comparing the holdings of a mutual fund with the holdings of each benchmark, we can connect each fund to one or multiple benchmarks and recover their complete investment style. By doing so, we define styles that are free from self-specified benchmark manipulation (Sensoy 2009).[7] Our matching algorithm and similarity-based weighted scheme to aggregate benchmarks constituents allow us to recover the full distribution of funds with respect to the size of their universe. *Large-universe* (LU) mutual funds and *Small-universe* (SU) mutual funds are located respectively above and below the median value. Third, Morningstar introduced to the public in 2016 a novel asset-weighted composite sustainable portfolio score[8] (Sustainability score) with a funds' coverage starting in 2012 (see, for instance, Candelon, Hasse & Lajaunie 2018, Ammann, Bauer, Fischer & Müller 2019 for recent studies using Morningstar's sustainable investment metrics).[9] We flag as *SRI* the top 10% and the rest as *Conventional*. We thus have four categories: SRI_{LU}, SRI_{SU}, $Conv_{LU}$, and $Conv_{SU}$.[10] Equipped with our sample of mutual funds segmented in subcategories, we apply standard regression to explain the Carhart (1997) four-factor alpha with respect to our categories and state-of-the-art controls (Cremers & Petajisto 2009). Eventually, as the performance of the funds could also be a driver of its migration toward the SRI strategy, we also test for the presence of endogeneity.

[6] Sectors (Kacperczyk et al. 2005), size, along with book-to-market ratio (Hoberg et al. 2017) are the most standard segmentation criteria. Morningstar also proposes categories based on geographical location (*e.g.*, World, Foreign, and Europe). These categories are not relevant in our case, as we focus on only US stocks. Among academic papers referring to the Morningstar style categories, see Brown & Goetzmann (1997) or Almazan et al. (2004).

[7] Overall, more than 87% of the funds' portfolio holdings in our sample are included in the Morningstar benchmarks.

[8] As explained above, SRI funds are characterised by a double investment objective, on the one hand taking into account ESG criteria, and on the other hand providing financial returns for their client. Yet, the degree with which each objective is pursued may vary between funds, thus we argue than focusing on Morningstar sustainable score offers a good opportunity to highlight truly SRI driven funds and test whether they are able provide financial returns on par with more conventional funds managers.

[9] Note that Ammann et al. (2019) use the Morningstar globes, which is a version of the score slightly different from ours.

[10] We alter our algorithm in different ways to test the robustness of our results in section 5.

Our study contributes to the literature on asset management in general and to the one on SRI in particular. Three features particularly stand out.

First, we develop a novel measure of investment universe based on investment styles that accommodates single and multi strategies. This measure enables to sort funds by the size of their universe. Unlike portfolio's size or portfolio's concentration for which we can find several measures in the literature (Kacperczyk et al. 2005, Brands et al. 2005, Fulkerson & Riley 2019), serious quantification of investment universe has been so far overlooked in academic studies to our knowledge. Our style-based approach is simple and flexible. It can be applied to any sort of funds for which we possess information on portfolio holdings.

Second, we obtain important insights from our results. As discussed in the literature, the effect of investment restrictions in financial performance is contrasted. On the one hand, it could worsen portfolio diversification. On the other hand, it helps managers better screen their assets, thus increasing the chance to size good opportunities and target companies with more sustainable profitability and better long-term prospects (Clark, Feiner & Viehs 2015, El Ghoul & Karoui 2017). One or the other effects might dominate, depending on a number of features among which the actual size of the universe—for example, with a very restricted universe, the adverse effects of poor diversification might dominate the screening effect. Separating SRI mutual funds and Conventional mutual funds by the size of their universe, we find that only SRI mutual funds with the smaller universe perform worse than Conventional funds. In accordance, our results show that, while the inclusion of socially responsible criteria into investment processes likely restricts the number of eligible assets with the risk of worsening portfolio diversification, such restrictions can be loosened by adopting a style benchmark strategy that increases the number of available assets (*e.g.*, multiple benchmark strategies or single benchmark with a large benchmark). This result is important, as its shows that SRI active investment managers would be well advised to consider a style that relies on a large range of assets in their portfolio management to minimize the side effect of restrictions imposed by SRI principals. From the end-user perspective, we provide empirical evidence to feed the debate on what risk-adjusted performance investors would require consistently in pursuit of ethical goals. These findings are robust to a series of robustness check.

Third, we propose an novel approach based on instrumental variable to explore the causal relationship between sustainability of mutual funds and their financial performance. As further developed below, studies explicitly tackling this issue are rather scarce. Our instruments pass both the usual relevance and exogeneity tests. Eventually, we find no evidence of endogeneity of SRI categories in our data.

The remaining paper is organized as follows. In section 2, we briefly summarize the literature. In section 3, we describe our methodology and data sources. Section 4 presents our results. Robustness analysis is performed in section 5. Section 6 concludes the study.

3.2 Literature Review

Our work follows three strands of the literature. In what follows, we review the studies that best relate to our research.

We first consider the broad literature on portfolio policy constraints (Koski & Pontiff 1999, Deli & Varma 2002, Almazan et al. 2004). Within this body of research, Almazan et al. (2004) explore the link between fund returns and a policy constraint index constructed as the equally weighted sum of restrictions regarding a fund's ability to use leverage (*e.g.*, borrowing of money, margin purchases, and short selling), its ability to use derivatives (*i.e.*, writing or investing in options on equities and writing or investing in stock index futures), and its ability to invest in illiquid assets (*i.e.*, investments in restricted securities). Based on the data of 9,525 US domestic equity funds from between 1994 and 2000, the authors find no differences in fund performance stemming from the level of investment policy restriction. Other contributions documenting alternative types of restrictions include Koski & Pontiff (1999) and Deli & Varma (2002) for the use of derivative securities or Clarke, De Silva & Thorley (2002) on position size and portfolio turnover effect. By focusing on SRI investment funds that attract an increasing amount of flows from investors while having to deal with restricted universe, we extend this line research. Our work also departs from existing contributions—We confront two types of restriction, namely, the one on socially responsible constrictions and the one on investment style, in order to assess whether the impact of one restriction can be alleviated by relaxing the other. To our knowledge, our study is the first to document whether the reduction in eligible assets from the adoption of a specific criteria or legal issues can be offset by the adoption of a style associated with a large universe.

Our study also contributes to the expending list of academic works on SRI mutual fund performance that test whether fund managers are "paying the price for ethics". Here, empirical evidence is mixed so far.[11] Some studies fail to report any statistical difference between risk-adjusted returns of the SRI and conventional mutual funds (*e.g.*, Bauer et al. 2005, Barnett & Salomon 2006, Gregory & Whittaker 2007, Meziani 2014, Dolvin, Fulkerson &

[11] Our main interest is to contrast the performance of high SRI funds with other funds conditionally on their universe. One line of enquiry in the SRI literature also explores the performance of "sin", that is, low SRI funds or stocks (see, for instance, Borgers, Derwall, Koedijk & ter Horst 2015, Richey 2016, Trinks & Scholtens 2017). For the sake of parsimony, we do not review this literature, as it fits less with our purpose.

Krukover 2019). Meanwhile, other contributions such as Renneboog et al. (2008*a*) find weak evidence that socially responsible funds under-perform compared with other funds, especially for the period between 1993 and 2003, based on a large set of funds across the world. Such under-performance is particularly strong in normal time, while SRIs may outperform in time of market crisis (Nofsinger & Varma 2014). Evidence of under-performance also exists for stocks (Ciciretti, Dalo & Dam 2019). More recently, El Ghoul & Karoui (2017) use data from the CRSP Mutual Fund Database along with the MSCI ESG KLD STATS database to construct a sample of mutual fund corporate social responsibility (CSR) scores by aggregating stock-level scores. Relying on a sample of 2,168 US equity funds between 2003 and 2011, they provide support for Renneboog et al. (2008*a*). That is, an increase in the level of CSR is negatively related to the fund's risk-adjusted performance. Our study extends these previous findings. First, we use Morningstar information on the sustainability score of mutual funds. By doing so, we can use an external score for the regression analysis. Second, our results tend to reconcile previous findings: We fail to observe significant difference between the SRI mutual funds taken as a whole and their Conventional counterparts. However, we show that SRI mutual funds do not form a homogeneous group, as those following a style associated to the smallest universe under-perform against their peers.

The final corpus of literature explores the causal relationship between sustainability of mutual funds and other mutual funds' features. Studies explicitly tackling this issue are rather scarce. For instance, Bauer et al. (2005), Renneboog et al. (2008*a*), and El Ghoul & Karoui (2017) remain silent about the identification issue stemming from the risk of reverse causality whereby financial performance influences the adoption of socially responsible restrictions. Renneboog et al. (2008*a*) use panel data at the monthly frequency to estimate the link between risk-adjusted returns and ethical funds. They disentangle the effect of various screening activities, such as the intensity of screening activity—for example, Number Sin Screens, Number Ethical Screens, and Number Social Screens—or the intent to influence corporate behavior through direct engagement. One dummy variable depicts whether a fund is an SRI or not. While the set of controls are time-varying and lagged by one period, the screening activity and SRI variables enter in the model as time-invariant over the 1991–2003 period. Proceeding this way likely mitigates the endogeneity problem. However, it also leads to erroneously portraying the funds' behavior, as screening activity generally evolves over time. By contrast, El Ghoul & Karoui (2017) exploit time variations in CSR and SRI practices. First, they show that the CSR score is explained by contemporaneous financial performance. Next, using a panel of yearly data, they lag all regressors, with CSR score and SRI dummy by one period, to explain the funds' alpha. Replacing a suspected simultaneously determined explanatory variable with its lagged value is a common practice in applied research. However, as discussed in the literature (see, Reed 2015), it might not be sufficient to rule out an endogeneity bias when the dependent variable displays serial correlation. Without further evidence on data properties, some of the existing results on the link between sustainable strategy and financial performance should, therefore, be taken with care. A more sophisticated approach to deal with this issue is proposed by Hartzmark & Sussman (2019), who explore the problem in depth. In 2016, sustainability ratings provided by Morningstar changed its manner to provide information on the sustainability score to the public opting for "globes" based on percentile. This shift, external to any fund-specific features, is considered to have dramatically improved the clarity of SRI ratings and their adoption by the market. In a quasi-experimental setting, Hartzmark & Sussman (2019) use this discontinuity as a shock to identify that the link between mutual funds ratings and funds' flows is caused by changes in rating.

Ideally, we would have an exogenous event in our study that induced a change in the SRI score, without directly affecting the risk-adjusted performance. We do not have such an event. However, we propose an original instrumental approach to test if our conclusions are robust to this caveat.

3.3 Data and Methodology

3.3.1 Data and sample

Quarterly data on mutual fund holdings are collected from the Morningstar Direct database. Over the past two decades, this source has been widely used in academic literature (see among others, Brown & Goetzmann 1997, Almazan et al. 2004, Del Guercio & Tkac 2008, Hartzmark & Sussman 2019). Morningstar is free from the survivor bias, and provides very comprehensive access to fund- and industry-level data. For the purpose of illustration, we can emphasize the disclosure of portfolio holdings, but also a large set of auxiliary information such as the Morningstar style benchmarks—the so-called "style box"—and Morningstar ratings, which are closely monitored by investors (Ammann et al. 2019, Hartzmark & Sussman 2019, Ben-David et al. 2019, Ciciretti et al. 2019). We restrict our analysis to active equity mutual funds domiciled in the US and investing in the US market[12]. Imposing this screening criteria keeps the sample relatively more homogeneous. It should also facilitate the recon-

[12] As done in Kacperczyk et al. (2005), Cremers & Petajisto (2009), Amihud & Goyenko (2013).

struction of the mutual funds' universe by limiting the set of available assets to US stocks[13]. We start with 3,172 US equity mutual funds[14] for the 2012Q4–2018Q4 period. We then apply standard filters (Kacperczyk et al. 2005, Ferreira et al. 2013, Kacperczyk et al. 2014) to meet our screening criteria. Specifically, we exclude fund of funds, index tracking funds, and equity mutual funds with less than two quarters of holdings observations. Furthermore, at each quarter—similarly to Cremers & Petajisto (2009)—equity mutual funds with less than US$ 1 million of asset under management and holding less than 60% of their portfolio in equity are dropped. Finally, we exclude funds for which no sustainability information is available (see section 3.3) and obtain a final sample of 2,039 distinct US equity mutual funds. We collect from the same source other usual characteristics of mutual funds, such as funds' gross and net monthly returns which will be used to compute the funds quarterly abnormal performances. We also retrieve their quarterly total net asset (TNA) and sustainability score. The sustainability mandate (*i.e.*, if the fund includes in its prospectus sustainability in its purpose) is also extracted. We further extract for each equity mutual fund the annual net expense ratio (NER) and turnover ratio[15] as well as the funds' age measured in years since inception of the oldest share. The missing values of NER are completed as in Elton et al. (2013) by subtracting the monthly gross return of our funds from their monthly net return, then retrieving their quarterly average. Regarding the turnover ratio, we complete each missing quarter by taking the percentage of a fund's TNA represented by the total number of assets purchased or sold between two periods, whichever is less (El Ghoul & Karoui 2017). With these data, we compute the fund flow, as in Cremers & Petajisto (2009), as the percentage growth in total TNA between two consecutive quarters, as well as their volatility over the previous 12 months.

3.3.2 Measuring mutual funds universe

A critical feature of the analysis pertains to the measurement of the mutual funds' universe. The universe of investment corresponds to the whole set of assets that each fund considers as candidates consistently with its general strategy when building and re-balancing its portfolio. Generally, it does not correspond to the entire set of tradable assets in the market (*i.e.*, "market portfolio"), but to a subset. Typically, it encompasses all of the securities in a particular asset class or of a fraction of the asset class based on additional segregation parameters. The

[13]Fund with less than 50% of their portfolio matching the entire stock universe of all US styles benchmarks are not considered

[14]When multiple shares are available, we only consider the "oldest share" identified by Morningstar to avoid redundancy, as in Porter & Trifts (2014)

[15]We then divide both ratios by four in order to obtain quarterly data.

resulting "restricted" universe defines the style of an institution, which plays a critical role in the mutual fund industry (Brown & Goetzmann 1997). As discussed in Almazan et al. (2004), for instance, part of the restrictions that face an investment fund capture the set of constraints that investors and managers consider as necessary to best define the fund's investment trademark. The style provides them with a simple tool to communicate their general strategy to their clients, eases performance comparisons with relevant peers, and helps concentrate their attention to a set of assets they know well. The style classification of equity is often defined along four dimensions (Brown & Goetzmann 1997, Kacperczyk et al. 2005, Hoberg et al. 2017): (i) size, (ii) book-to-market ratio, (iii) industry, and (iv) geographic location. Typically, stocks from firms of similar size are placed in the same style group. The same applies for other common segmentation criteria, such as fundamental characteristics (*e.g.*, value vs. growth stocks) or industries (*e.g.*, financial vs. industrial). Accordingly, the whole set of equities in the market can be allocated into relatively homogeneous groups of assets, namely, benchmarks wherein each fund will consider making its selection to build its portfolio. Investments into benchmarks are not exclusive. Some managers may opt for a single-benchmark style (*e.g.*, value), while other rely on multiple-benchmark style (*e.g.*, large-cap/value or small-cap/value), spreading their investment across various groups to increase their diversification, for instance.

Against this background, we propose an original approach to recover the universe of each fund included in our sample across time. To do so, we exploit granular information about benchmark composition along with portfolio holdings of funds. The underlying idea of our approach is to associate each fund with one or several benchmarks, depending on their actual portfolio holdings. We call them the "associated benchmarks." Then, we aggregate the number of assets included in all associated benchmarks. The weights used in the aggregation step depend on how much important a benchmark is in the fund's portfolio. Lastly, we ultimately aim to create categories of large versus small universes for which recovering a ranking of funds is sufficient. For instance, we do not need to know that fund A and fund B do have a universe of 400 assets and 300 assets, respectively, for instance, but just that A possesses a larger universe than B. In accordance, we actually do not have to estimate the exact size of their universe, but only recover proxies good enough to preserve the relative positioning of each fund in the ranking. As we cannot rule out the presence of noise in our metric, we propose to test the robustness of our findings to alternative approaches for the classification.

Formally, at each quarter, we measure the investment universe for each fund as the total number of stocks held by the style benchmarks associated with a fund, scaled by the overlap between the fund portfolio and the benchmarks, and label this variable $IU\ Score_{f,t}$. Thus,

$$IU\ Score_{f,t} = \sum_{b=1}^{B} (w_{f,b,t}) * N_{b,t} \tag{3.1}$$

where N[16] is the number of shares held by benchmark b at time t and $w_{f,b,t}$ is the weight attached to benchmark b for fund f at time t (see below).

To compute our measure, we need to break the entire set of tradeable equities into homogeneous groups (*i.e.*, benchmarks, b) and quantify them (*i.e.*, count the number of available assets, $N_{b,t}$). A natural and simple way to assign individual stock to meaningful style benchmarks is to use Morningstar-style categories. Since almost three decades, these categories have been widely monitored by the industry Haslem (2009, 2017).

Morningstar categories emerged in the early nineties amid suspicion of strategic manipulation of self-declared styles by mutual funds. To provide a more accurate assessment of actual investment style, the Chicago-based firm developed investment categories with a list of their components (*i.e.*, associated assets) to be matched with mutual funds' holdings in order to recover the actual investment strategy of mutual funds as opposed to the one stated in their prospectus. The so-called "equity style box," for instance, consists of a 3x3 grid based on the market capitalization as well as growth and value factors, which allows traded securities to be grouped into nine "investment-style" categories: large cap value, large cap blend, large cap growth, mid-cap value, mid-cap blend, mid-cap growth, of value/blend/growth and small/mid/large-cap. These categories can be applied to both stocks and funds.[17] In addition to these nine categories, Morningstar also defines categories by sectors: Communications, Consumer Cyclical, Consumer Defensive, Energy Limited Partnership, Energy, Precious Metals, Financial, Healthcare, Industrials, Infrastructure, Miscellaneous Sector, Natural Resources, Real Estate, Technology, and Utilities. We do not consider geographical dimension, as our sample is restricted to US equities. In total, we consider 26 Morningstar categories or benchmarks b. These benchmarks embed 3,826 stocks on average per period that is nearly 100% of the U.S. investable equity market. The style attributes of individual stocks are then used to identify the style classification of mutual funds and eventually recover

[16] As explained below, we apply a small correction to our data. As benchmarks are not exclusive and could have some assets in common, we make sure that an asset appears only one time in a fund's universe.

[17] Note that growth and value categories are common to both stocks and funds. Growth funds are deemed so if the investment is in growth stocks. For the blend category, the situation is slightly different. Stocks wherein neither the value nor the growth characteristics dominate are labeled core. For funds, the category next to value and growth is the blend style. It can be a mixture of growth, value stocks, as well as core stocks. However, in practice, as stated in Morningstar documentation, it "mostly [constituted by] core stocks" www.morningstar.be/be/news/article.aspx). For the sake of simplicity, we keep the term "blend" to characterize both stocks and funds.

their universe. To match each fund with its associated benchmarks, we measure the overlap of the funds' portfolio relative to each benchmark. To do so, we use an overlap's measure based on the Manhattan distance, following Cremers & Petajisto (2009), for constructing their Active Share (AS) score[18]. Formally, we measure the proximity of each fund "f" with any benchmark "b" at each quarter "t" as follows[19]:

$$w_{f,b,t} = (1 - AS_{f,b,t}) = (1 - \frac{1}{2}\sum_{i=1}^{M} |\omega_{f,i,t} - \omega_{b,i,t}|) \tag{3.2}$$

where M is the sum of stocks in either the portfolio of f or the benchmark b at time t. $\omega_{f,i,t}$ is the share of stock i in the portfolio of f and $\omega_{b,i,t}$ is its share in the benchmark b. This measure ranges from 0% to 100% (if no short-selling is allowed), where a score of 100% indicates full overlap between the fund and its benchmark portfolio. We proceed by sorting the benchmarks for each fund by the level of overlap—from the closer benchmark to the more distant. We pick the first one and count its number of assets $N_{1,t}$. To identify the number of assets increasing the size of the universe in the second benchmark, we have to monitor any potential redundancy of assets across benchmarks (*i.e.*, assets appearing in multiple benchmarks). We remove assets from the others benchmarks which were common to the first one and count the number of remaining assets $N_{2,t}$ included in the second match. We repeat this procedure as we go down in the list until all benchmarks have been covered. If a benchmark has zero stocks left in common with the fund (after the iterative cleaning), it is dropped. All remaining benchmarks are deemed "associated benchmarks." They all define the style of the fund at different degrees. As a consequence, all their assets can be considered consistent with the fund's strategy and, as such, part of its investment universe. All associated benchmarks are probably not of the same importance for a fund and so should be the weights attached to the benchmarks' constituents when computing the universe. For instance, a fund could hold 99% in a small benchmark and only a 1% in a large benchmark. Weighting assets populating both benchmark as the same could inflate its universe. We control for the importance of a benchmark in the aggregation step by weighting each asset with our overlap measure $w_{f,b,t}$. Note that the weighting step makes it difficult to interpret the value of the score, which should be viewed as a metric to sort mutual funds by the size of their universe.

[18] AS is a measure of dissimilarity between a fund and a benchmark

[19] In our study, we focus on the equity portfolio of mutual-funds; if the portfolio is not made of a 100% of equity, we re-scale its weights to sum to 100%

3.3.3 Uncovering SRI-driven funds

We use the Morningstar Sustainability Rating as our source of data on sustainability. The rating enables us to assess a fund's level of sustainability and, thus, compare it to others. The score provided by Morningstar is holdings-based, that is, it reflects the sustainability of the firms held in a mutual fund's portfolio. Specifically, the rating stems from the aggregation of the normalized firms' ESG scores and controversy score. The former reflects to what extend a firm successfully manages ESG risks and opportunities relative to their industry peers. It ranges from 0 to 100. The later tracks involvement in ESG-related incidents that may negatively affect the environment or society, and its related risk for the company itself. The controversy score ranges from 0 to 100. Both scores are provided by Sustainalytics, a global leader in sustainability research and analysis in the industry. With firm-level data, Morningstar uses the asset-weighted sum of the difference between the *ESG* score and the *controversy* score to compute the *sustainability score*:

$$Sustainability\ score = \sum_{i=1}^{n} \omega_i (ESG_i - Controversy_i) \tag{3.3}$$

where ESG_i and $Controversy_i$ are the industry normalized firms' ESG score and controversy score, respectively, and n is the number of assets composing the portfolio[20] and ω_i its share. The score is updated monthly based on the most recent company data from Sustainalytics.

In our context, the Morningstar Sustainability Rating has a number of advantages. First, Morningstar emerged rapidly as a recognized source of information in the literature on mutual funds' sustainable investment practices (Ammann et al. 2019, Hartzmark & Sussman 2019, Ceccarelli, Ramelli & Wagner 2019). Second, it covers a very large part of the market over a reasonably long period of time. In general, most of the funds are rated by the end of 2012. More than 90% of the US equity mutual funds included in our sample did have a sustainability rating from Morningstar. Third, the construction of the score is fully transparent. Fourth, it provides an external measure of sustainability that could easily and directly be used by many researchers, thereby avoiding the need to develop a new measure, as in El Ghoul & Karoui (2017). Based on the sustainability score, we break our sample of mutual funds into categories. To do so, we rank all mutual funds at each quarter according to their *score*. The

[20] A fund receives either an ESG or controversy score, if this information is present for at least 50% of a portfolio's assets under management. The scored assets' weights are then re-scaled to 100% (for further details on the computation of the score, see https://s21.q4cdn.com/198919461/files/doc_downloads/press_kits/2016/Morningstar-Sustainability-Rating-Methodology.pdf

mutual funds in the top 10^{th} decile are labeled "*SRI*" and the rest "*Conventional*." The 10% threshold is consistent with the proportion of self-declared "socially conscious" mutual funds in their prospectus.[21] Eventually, we combine this information with the one on the size of the universe to create our main variables of interest. Taking each group separately, we rank them with respect to their *IU Score*. Large-universe (LU) mutual funds and Small-universe (SU) mutual funds are located respectively above and below the median values. We then have four categories: *SRI_LU*, *SRI_SU*, *Conv_LU*, and *Conv_SU*.

3.3.4 Modeling (un)conditional effect of SRIs on mutual funds' performance

Our framework enables us to formulate two testable hypotheses:

- H1: Are SRI principals detrimental to mutual funds' financial performance?
- H2: Are restrictions in an investment universe imposed by sustainability considerations compensated for in other areas by adopting styles associated with large universes?

To do so, we estimate a panel regression forecasting the four-factor abnormal returns $\alpha_{i,t+1}$ based on the following model[22]:

$$\begin{aligned}\alpha_{f,t+1} \quad = \quad & \beta_0 + \beta_{LU} * SRI_{LU_{f,t}} + \beta_{SU} * SRI_{SU_{f,t}} + \delta_{SU} * Conv_{LU_{f,t}} + \delta_{LU} * Conv_{SU_{f,t}} \\ & + \textstyle\sum_{i=1} \beta_i * Controls_{i,f,t} + Q_t + S_f + \epsilon_{f,t+1}\end{aligned} \tag{3.4}$$

where $SRI_{LU_{f,t}}$ is a dummy variable that takes on the value 1 if a fund f at time t is SRI-driven and with a *IU Score* above the median, and 0 otherwise; SRI_{SU} is a dummy variable that takes on the value 1 if a fund is SRI-driven and with a *IU Score* below the median and 0 otherwise; $Conv_{LU}$ is a dummy variable that takes on the value 1 if the funds is not SRI-driven and with a *IU Score* above the median, and 0 otherwise; $Conv_{SU}$ is a dummy

[21]We count 188 self-declared socially conscious funds in our sample to be compared to a total of 2,039 mutual funds, that is, $\sim 10.8\%$

[22]We compute the funds' risk-adjusted returns by estimating the traditional four-factor alpha (Carhart 1997) with rolling windows of 24 months (complete return information over the window is required) moving by one quarter. The monthly US factors and risk-free rate are retrieved from the Kenneth French online library (available at: https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html).Using the four-factor alpha allows to ensure that the performance of SRI funds are not driven by their exposure to specific equity characteristics (e.g. growth and large capitalisation styles. For comparison purpose, we also implemented the 5-factor alpha (Fama & French 2010) as in Joliet & Titova (2018) and no substantial differences were observed (results are available upon request)

variable that takes on the value 1 if the funds is not SRI-driven and with a *IU Score* below the median, and 0 otherwise.

The first hypothesis, $H1$, corresponds to the unconditional effect of the SRI on mutual funds' performance. To test it, we impose $\beta_{LU} = \beta_{SU}$, that is, the influence of the SRI on financial performance is homogeneous across SRI mutual funds. In addition, we set $\delta_{LU} = \delta_{SU} = 0$ to compare it to all Conventional mutual funds and avoid perfect collinearity.

The second hypothesis, $H2$, corresponds to the conditional effect of the SRI on mutual funds' performance with respect to the size of the universe. To test it, we relax the homogeneity assumption and freely estimate β_{LU} and β_{SU}. To avoid perfect collinearity, we set one parameter among $\delta_{LU}, \delta_{SU}, \beta_{LU}$, and β_{SU} equal to zero. The selected parameter provides the base effect against which the three other coefficients should be compared.
Our goal is to identify the marginal impact of the adoption of SRI principals on financial performance along with the role that style benchmarks can play in mitigating the resulting restrictions on eligible assets. Both features could be correlated with others traits of mutual funds, such has their age or their size, which have been shown to be significant determinants of performance in the literature (see, for instance, Cremers & Petajisto 2009, Ferreira et al. 2013, Amihud & Goyenko 2013). To isolate our effects, we add a large set of controls. Following the literature Cremers & Petajisto 2009, Amihud & Goyenko 2013, we include the funds size, computed as the natural logarithm of their TNA (*LN(TNA)*), as well as its squared value ($LN(TNA)^2$), the natural logarithm of the funds age (*LN(Age)*), and the natural logarithm of the number of assets composing the portfolio (*LN(Assets)*). We further add the funds' flows (*Flow*) as well as their volatility over the previous 12 months (*Flow-sd*), turnover ratio (*Turnover)*, (*NER*), and the R-square (R^2) of the four-factor performance regression as a stock selectivity proxy (see Amihud & Goyenko 2013). Our final controls is the funds' volatility return (*Vol*) computed as the standard deviation of monthly returns over the previous 24 months. Eventually, in line with El Ghoul & Karoui (2017), we add time (Q_t) as well as style (S_f) fixed effects. As usually done in the literature, we lag all controls by one period and winsorize the continuous variable at the 1% level[23].

[23] To ease the interpretation of our results, we standardized each continuous variable (zero mean and unit standard deviation)

3.4 Results

3.4.1 Descriptive statistics

Table 3.1 reports the summary statistics of our main variables. The mutual funds included in our sample are, on average, worth US$ 1,71 billion and are 17 years old. The average quarterly net flows is negative (-0,12%) with a standard deviation of 12.38%, reflecting a slowdown in this segment of the industry over our sample period, but also marked heterogeneity. The mean of the quarterly net abnormal performance is negative (-0,49%), consistent with the literature on active management having documented its inability to create value for investors on average. The average quarterly turnover is around 15%, while their annual NER is about 1%. The fund SRI score ranges from 28.01 to 58.33, with an average of ~ 45.88. The typical active manager holds 188 stocks over four main associated benchmarks but invests more than 40% in its main benchmark (best match). The R^2 of the four-factor alpha regression is at the median equal to 93.28% indicating that most of the average fund's returns variability is explained by the Carhart (1997) four-factor model. Turning to the investment universe score, we observe an average value of 228 with a minimum value of 31 and a maximum value of 477. The unweighted measure of investment universe[24] does not account for fact that all associated benchmarks are not of equal importance for a fund. However, its value is easier to interpret. The mean of stocks consistent with funds strategies is of 1,556 with a maximum of 3,201 stocks.

[24] We select the benchmark with which the portfolio has at least 5% in common and sum their total number of unique assets

Table 3.1: Summary statistics

	Mean	Median	Sd	Min	Max
Age (Years)	17.40	15.75	13.44	0.50	70.08
Alpha 4-factor (% Quarterly)	-0.49	-0.44	0.60	-4.53	3.057
Flow (% TNA)	-0.12	-0.63	12.38	-12.48	29.21
Flowsd	3.72	1.56	7.75	0.16	63.84
#Assets	118	70	180	20	1466
NER (%)	1.00	0.98	0.3440	0.20	2.42
$\mathbf{R^2}$ (%)	89.41	93.28	12.41	28.43	99.38
Styles best match (%)	42.07	42.13	13.9	11.49	78.76
Styles covered	3.7	4	1.65	7	1
Sustainability score	45.88	46.17	2.97	28.01	58.33
TNA ($ bn)	1.71	0.42	3.55	0.002	21.85
Turnover (%)	14.67	11.50	12.08	0.50	77.25
IU Score	228	218	107	31	477
IU Score (unweighted)	1584	1526	727	161	3201
Vol (%)	3.53	3.41	0.89	2.03	6.96

Note: Table 3.1 reports the descriptive statistics of our main variables over the 2012Q4 to 2018Q4 period. More precisely it reports the mean, median, standard deviation minimum and maximum of the funds's age expressed in years, the risk adjusted performance (*alpha 4-factor*) over the quarter and its associated R^2, the TNA in billion dollar (*TNA*), the flows and flows' volatility over the last 12 months, the % of the funds belonging to its best match style, the number of styles covering 80% of the portfolio and accounting for at least 5% of the portfolio individually (*Styles covered*),the number of assets composing the portfolio (*Assets*), the (*NER*) as well as the turnover ratio, the volatility of the funds returns over the past 24 months, the sustainability score and the funds weighted and unweighted investment universe score (*IU Score*).

In Table 3.2, we report summary statistics regarding benchmarks. We observe that benchmarks are constituted, on average, of 126 stocks, for the smallest, and 549 for the largest. Next, we look at the allocation of funds across benchmarks. We consider each time the main benchmark of each fund that is the one with the greatest match. The average number of funds is around 185 for the equity style box styles and 22 for US sectors. We can notice that SRI and Conventional funds are distributed rather similarly across benchmarks (with a correlation of 75.95%) with, for both, a majority of investments made in large caps.

Table 3.2: Mutual funds' styles descriptive statistics

Name	**Type**	**#Funds**	**#Assets**	**%Conventional**	**%SRI**
Communication	US Sector equity	4	301	0,28	0.00
Consumer Cyclical	US Sector equity	8	292	0,55	0,72
Consumer Defensive	US Sector equity	10	248	0,52	1,92
Energy Limited Partnership	US Sector equity	13	126	0,84	0,99
Equity Energy	US Sector equity	21	309	1,40	1,38
Equity Precious Metals	US Sector equity	3	194	0,11	0,63
Financial	US sector equity	29	349	1,99	1,58
Health	US sector equity	31	260	1,61	6,23
Industrials	US sector equity	6	321	0,28	1,27
Infrastructure	US sector equity	0	360	0.00	0,00
Large Blend	US Equity	106	267	6,45	12,34
Large Growth	US Equity	308	325	21,74	11,63
Large Value	US Equity	394	343	25,59	33,39
Mid-Cap Blend	US Equity	53	457	3,88	0,79
Mid-Cap Growth	US Equity	127	419	9,40	0,88
Mid-Cap Value	US Equity	100	484	7,37	1,10
Miscellaneous Sector	US Sector equity	3	549	0,14	0,76
Natural Resources	US sector equity	8	524	0,54	0,63
Real Estate	US sector equity	54	197	2,18	15,48
Small Blend	US Equity	46	408	3,46	0.00
Small Growth	US Equity	64	447	4,79	0.00
Small Value	US Equity	44	456	3,25	0.63
Technology	US sector equity	57	455	3,44	6,96
Utilities	US sector equity	3	334	0,13	0.63
				— Correlation : 75.95%	

Note: Table 3.2 reports the average number of funds associated to each styles (according to their best match) at each quarter throughout our sample, as well as the styles' type and the style average number of assets. %Conventional and %SRI represent the average proportion of each type of funds represented in each style. The styles are determined by Morningstar to categorize funds with similar holdings together and create meaningful categories for the investors. Each style as an associated benchmark for which we have complete holdings information. We have a total of 24 US styles, of which 15 are based on sectors and 9 on stocks capitalization. Infrastructure is included in our calculations, even though no funds in our sample uses it as their main style, they might still allocate part of their remaining wealth to it.

Table 3.3 reports additional features. We observe that SRI funds tend to be slightly younger and larger than their Conventional counterparts. They hold significantly less stocks and focus on a lower number of benchmarks. They exhibit more stable flows and a higher degree of stocks selectivity, as expressed by their lower R^2. Conversely, Conventional funds display, on average, a much higher turnover ratio. Eventually, without controlling for additional features, SRI funds exhibit higher performance than Conventional ones.

Table 3.3: Characteristics of SRI vs Conventional funds

	Conv	**SRI**	**SRI-Conv**	**t-statistic**
Age (Years)	17.43	16.99	-0.44*	-1.7
Alpha 4-factor (% Quarterly)	-0.51	-0.32	0.19**	2.2
Flow (% TNA)	-0.30	0.37	0.67	1.4
Flowsd	0.037	0.036	0.001*	-1.7
#Assets	124	65	-59***	-23
NER (%)	0.9997	1.0025	0.0028	0.81
R^2 (%)	90.15	82.65	-7.5***	-8
Styles Best match (%)	41.47	47.48	6.01***	6.7
Styles covered	3.91	3.69	-0.022***	-3.4
Sustainability score	45.48	49.72	4.14***	10
TNA ($ bn)	1.67	2.10	0.43***	3.1
Turnover (%)	15.01	10.97	-4.04***	-10
IU Score	235	204	-31***	-7.1
IU Score (unweighted)	1584	1316	-268***	-7.3
Vol (%)	3.56	3.33	-0.23	-1.2

Note: Table 3.3 reports the average level of our main variable relatively to their main category of funds (SRI vs Conventional). Variables are defined as in Table 2.2. Each quarter we flag the funds having the best sustainability score (top 10%) as SRI funds and the rest as Conventional (Conv), then we aggregate at each quarter the fund's variable for each category which we display in the first and second column. The third column displays the difference in means between the variable stemming from SRI funds and Conventional ones. The last column reports the t-statistics for the differences in means.

3.4.2 Unconditional effect of SRI on mutual funds' performance

Before gauging the relationship between funds' performance and their sustainability score, Table 3.4 reports the coefficient of the quarterly panel regression of funds' risk adjusted performance on state-of-the-art characteristics. In the table, we indicate both the value of the coefficients and the t-statistics based on standard errors clustered at the fund level. In its lower part, we include whether or not we use time and/or category fixed effects, the number of funds, and the number of observations. Overall, the results on the contribution of our controls are consistent with the literature. The turnover and (NER) are significant negative drivers of future performance, as in Carhart (1997), which may imply that excessive turnover increases transaction cost, which, in turn, lowers the value for investors. Funds managers charging more on the claim that they have superior skills actually decrease the final net returns for their investors. As Cremers & Petajisto (2009) and Amihud & Goyenko (2013), we find a weak negative link between the fund's age (*Ln(Age)*)[25] and its abnormal returns. No significant effect of the funds' size can be identified.

In line with Amihud & Goyenko (2013), who states that the R^2 stemming from a multi-factor performance regression can be used as a manager "stock selectivity" proxy[26], our results show that the higher the manager's selectivity (the lower the R^2), the higher the fund subsequent performance. Turning to the positive sign of flows ($Flow$), the result supports the findings of Gruber (1996) and the "smart money" effect, thereby investors are able to uncover skilled managers and invest in their funds. Yet, excessive volatility in the flows ($FlowSd$) is detrimental, suggesting that funds having more stable sources of incomes tend to fare better. We find that excessive volatility negatively impacts future abnormal returns. Eventually, the natural logarithm of the number of assets composing the portfolio has no explanatory power over the four-factor alpha.

[25] Amihud & Goyenko (2013) find no significant effect of age on performance

[26] The main intuition is that if the R^2 is high then the manager is mostly passively tracking the market

Table 3.4: Baseline model of mutual funds performance

	Alpha_{t+1}
LN(Assets)	0.016
	(1.57)
Flow	0.211***
	(20.08)
Flowsd	-0.035***
	(-3.82)
LN(Age)	-0.023*
	(-1.81)
LN(TNA)	0.125
	(0.75)
LN(TNA)^2	-0.039
	(-0.25)
NER	-0.111***
	(-9.20)
R2	-0.112***
	(-5.90)
Turnover	-0.077***
	(-6.62)
Vol	-0.419***
	(-20.06)
Time fixed effects	Yes
Category fixed effects	Yes
Number of Funds	2039
Number of observations	35455

t-statistics in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Note: Table 3.4 reports the baseline panel regression of the performance measure (alpha 4-factor) on lagged control variables. *LN(Assets)* represents the natural logarithm of the number of unique positions composing the portfolio. *Flow* designate the net fund level flow in % of the lagged value of TNA and *Flowsd* the flow volatility over the previous 12 months. *LN(TNA)* is the natural logarithm of the fund TNA and *LN(TNA)*^{2} its squared value. *LN(Age)* is the natural logarithm of the fund age expressed in years. *NER* and *Turnover* represent respectively the prospectus Net expense and turnover ratios. Finally, R^2 is the R-squared of the 4-factor regression and *Vol* the volatility of the funds returns over the last 24 months. We include time and style fixed effect and t-statistics are based on standard errors clustered at the fund level.

In Table 3.5, we include our SRI variable in the model. Its coefficient is negative, but not statistically significant. Consistent with the majority of existing academic studies, we cannot reject the null hypothesis that SRI funds' abnormal returns, once we control for other characteristics, are not different from their conventional counterparts.[27]

Table 3.5: Unconditional effect of SRI on mutual funds performance

	Alpha_{t+1}
SRI	-.040
	(-1.44)
LN(Assets)	0.017
	(1.45)
Flow	0.211***
	(20.08)
Flowsd	-0.035***
	(-3.82)
LN(Age)	-0.023*
	(-1.80)
LN(TNA)	0.124
	(0.75)
LN(TNA)^2	-0.039
	(-0.25)
NER	-0.111***
	(-9.20)
R2	-0.112***
	(-5.90)
Turnover	-0.077***
	(-6.62)
Vol	-0.420***
	(-20.06)***
Time fixed effects	Yes
Category fixed effects	Yes
Number of observations	2039
Number of observations	35455

t-statistics in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Note: Table 3.5 reports the panel regression of the funds' performance on controls variables (defined as in table 2.5) and our SRI dummies. The SRI dummy takes on the value 1 if the funds' portfolio is in the top 10% of all funds' portfolio sustainability score in quarter q, 0 otherwise. We include time and style fixed effect and *t*-statistics are based on standard errors clustered at the fund level.

[27] Statman (2000), Bauer et al. (2005), Renneboog et al. (2008*b*), Derwall et al. (2011), Meziani (2014)

3.4.3 Conditional effect of SRI on mutual funds' performance

So far, the sample of SRI mutual funds has been assumed to be homogeneous. We relax this assumption by segmenting both SRI mutual funds and Conventional mutual funds by the size of their universe. Table 3.6 reports the results. In Column I, the category SRI_{SU} is dropped. In accordance, the coefficients of SRI_{LU}, $Conv_{SU}$, and $Conv_{LU}$ are expressed with respect to SRI_{SU}. All three categories are shown to have a significant positive impact on future performance. If we consider SRI_{LU} and $Conv_{LU}$, the coefficients are statistically significant at the 1% level. $Conv_{SU}$ is significant at the 5% level. The sign is positive, that is, all three categories—including SRI mutual funds with larger universe—perform better than the SRIs with small universes. The predicted difference in abnormal return between SRI mutual funds with large universe and small universe is about 0.16%. Turning to Column II, we observe that SRI mutual funds with large universe not only outperform those with tighter universe from the same category, but also Conventional funds with small universes (albeit weakly at the 10% level). The difference is not statistically significant with $Conv_{LU}$. Column III and IV reports the results when respectively $Conv_{SU}$ and $Conv_{LU}$ are used as base effect. The size of the universe is not critical among Conventional funds. While funds broadening their universe have higher average performance, the results are not statistically significant.

Overall, these results do show that the sample of SRI mutual funds is not homogeneous. Those with the smaller universe under-perform significantly against Conventional mutual funds, even though we previously failed to detect any difference when a sustainable investment strategy was applied. This finding supports our second hypothesis whereby ethical restrictions are not necessarily detrimental to financial performance as long as the managers are able to expand their universe in other areas (forming their portfolio by choosing a combination of benchmarks giving access to a large pool of stocks).

3.5 Extensions and Robustness

How robust are these findings? We conduct several robustness tests on the benchmark model. In this section, we closely scrutinize several important caveats and issues related to our empirical strategy. We deal with potential endogeneity issue before pivoting toward alternative definitions for creating our main categories: (i) SRI versus Conventional mutual funds as well as (ii) mutual funds with large and small universes.

Table 3.6: Investment universe and conditional effect of SRI on mutual funds performance

	Alpha_{t+1} I	Alpha_{t+1} II	Alpha_{t+1} III	Alpha_{t+1} IV
SRI_{SU}		-0.160***	-0.093**	-0.126***
		(-3.37)	(-2.29)	(-3.09)
SRI_{LU}	0.160***		0.067*	0.034
	(3.37)		(1.80)	(1.04)
Conv_{SU}	0.093**	-0.067*		-0.033
	(2.29)	(-1.80)		(-1.42)
Conv_{LU}	0.126***	-0.034	0.033	
	(3.09)	(-1.04)	(1.42)	
Controls	Yes	Yes	Yes	Yes
Time fixed effect	Yes	Yes	Yes	Yes
Style fixed effect	Yes	Yes	Yes	Yes
Number of Funds	2039	2039	2039	2039
Number of observations	35455	35455	35455	35455

t-statistics in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Note: Table 3.6 reports the panel regressions of the funds' performance on controls variables (defined as in table 2.5) and category dummies. SRI_{SU} takes on the value 1 if the fund's portfolio sustainability score is in the top 10% and if its universe (*IU Score*) is in the top 50%, conversely SRI_{LU} is in the bottom 50% of our *IU Score*. $Conventional_{SU}$ and $Conventional_{LU}$ follow the same logic but are in the 9 first deciles of the portfolio sustainability score. We include time and style fixed effect and t-statistics are based on standard errors clustered at the fund level.

Though the results in the previous section are suggestive, they cannot rule out the possibility that financial performance influences the adoption of the SRI strategy. In accordance, the first potential caveat being examined is whether the sustainable score may be endogenous. Such endogeneity is often disregarded in the literature. However, there are various reasons why we might suspect the adoption of high SRI standards to result from funds' performance. If true, it would lead to double causality and the ordinary least squares estimates reported in the previous section would be biased. For instance, while migration toward a more sustainable investment strategy could be fruitful in the long term because of being better suited with regulatory or customers future requirements, it can be costly in the short term. Such cost stems from the lack of knowledge or expertise of fund managers when they start implementing new strategies. To deal with this transition period, mutual funds may choose to do so when their financial and reputational capital is high. In this case, their performance would explain their score. An alternative mechanism that could also lead to reverse causality is, contrariwise to the previous one, that "losers" in the industry aim to change their policy and specifically move toward one for which the financial performance could be viewed as less important (Riedl & Smeets 2017). To confirm that previous estimates are unbiased, we formally test the presence of endogeneity. As usual, the cornerstone of the testing strategy lies in the choice of the instrument. Compelling instruments should satisfy both the exogeneity and the validity condition. As often in economics and finance, the literature on sustainable investment in the mutual funds industry provides little guidance for the choice of an instrument. The strategy we propose here is to rely on an instrument inspired from the literature on peers' effects. The intuition is that choosing to adopt sustainable criteria for investing could, to some extent, depend for a given fund on the behavior of its peers. Typically, if a group of mutual funds (*i.e.*, peers) are set in competition, some of them improving their sustainable score could lead others to align. Thus, considering the average score of the peers as an instrument directly is an option. However, as all funds among a peers' group might respond to each other, the score the of peers might also depend on the one we aim to instrument. Hence, as known in the literature on social interactions, group behavior is suspected to be endogenous. An alternative is to use exogeneous group characteristics. Bearing this in mind, we propose the age of peers' funds as a main instrument. We suppose that such variable influences the score of the peers, without being affected by neither the score of the instrumented fund nor its performance. Peers group are identified as funds following the same main style (*Best match*). For a given quarter, the value of the instrument for a fund i is the leave-one-out mean (*i.e.*, mean of the group excluding the fund in question) of its peers in the previous year. Table 3.7 reports the main statistics of our testing procedure. We find that peers' age has a negative and significant effect on mutual fund's sustainable score. The F-stat is equal to 15.98. Based on Staiger and Stock's rule of thumb critical value of 10, our instrument can be considered relevant. Given the nature of the instrument, we are confident about its exogeneity. Still, we also propose a formal test. The exogeneity assumption requires a second instrument to be assessed. We use a normalized Herfindahl–Hirschman index (HHI)[28]of portfolio weights as a second instrument in order to perform the Sargan–Hansen overidentification test. Such a measure enables to focus on the distribution of the weights attached to holdings rather than the overall level of concentration. Results from Sargan–Hansen test confirm the exogeneity of our instruments (see Table 3.7). Equipped with our instruments, we conduct an Hausman specification test to detect endogeneity in the sustainable score. The null hypothesis of no endogeneity could not be rejected at any usual significance levels whether we consider *peers age* as a standalone instrument or in combination with *HHI*.

[28] $HHI = \frac{\sum_{i=1}^{N} \omega_i^2 - 1/N}{1-1/N}$, where ω_i is a fund's portfolio weights and N is the number of unique positions

Table 3.7: Endogeneity test

Instruments	**Peer-age**		**HHI**		**Peer-age + HHI**	
	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value
First stage						
F-statistics of excluded instruments	15.98	0.00	4.84	0.028	10.66	0.00
Second stage						
Sargan-Hansen test	/	/	/	/	0.869	0.35
Hausman test	0.60	0.44	2.68	0.102	2.308	0.129

Note: Table 3.7 reports the main statistics of the instrumental variable regression. Peer-age refers to the average age of a fund's peers lagged by one year, HHI represents the normalized Herfindahl-Hirschman index of each fund's portfolio weights lagged by one year. We perform a two stage least square regression and report the results for the F-test of excluded instrument, the Sargan-Hansen test of over-identification and finally the Hausman test of endogeneity

As a next step, we use alternative category definitions to check how sensitive the results are to such changes. First, we strengthen our criteria for selecting SRI mutual funds and consider only those displaying time consistency in their strategy. To do so, we construct a new variable SRI_alt1 taking the value of 1 if a mutual fund is included in the top 10% at least 60% of the time during the sample period, and 0 otherwise. The resulting SRI_alt1 is time invariant as in Renneboog et al. (2008*b*). Table 3.8 confirms the main findings of the benchmark results, as SRI mutual funds with the smaller universe (SRI_alt1_{SU}) is the only category under-performing all the others. In this model, the SRI funds with the largest universe (SRI_alt1_{LU}) no longer outperform the conventional ones with the smallest universe ($Conv\_alt1_{SU}$).

Table 3.8: Conditional effect of SRI on mutual funds performance
Alternative SRI category: Persistent SRI ranking

	Alpha_{t+1}	Alpha_{t+1}	Alpha_{t+1}	Alpha_{t+1}
	I	II	III	IV
SRI_alt1_{SU}		-0.193**	-0.116**	-0.158***
		(-2.44)	(-2.13)	(-2.91)
SRI_alt1_{LU}	0.193**		0.076	0.035
	(2.44)		(1.09)	(0.52)
Conv_alt1_{SU}	0.116**	-0.076		-0.042*
	(2.13)	(-1.09)		(-1.86)
Conv_alt1_{LU}	0.158***	-0.035	0.042*	
	(2.91)	(-0.52)	(1.86)	
Controls	Yes	Yes	Yes	Yes
Time fixed effect	Yes	Yes	Yes	Yes
Style fixed effect	Yes	Yes	Yes	Yes
Number of Funds	2039	2039	2039	2039
Number of observations	35455	35455	35455	35455

t-statistics in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Note: Table 3.8 reports an alternative setting for Table 3.6 where SRI funds (SRI_alt1) are the one whose observation are at least 60% of the time in the top 10% of the portfolio sustainability score. We include time and style fixed effect and t-statistics are based on standard errors clustered at the fund level.

Second, we use a threshold of 20% rather than 10% to define SRI mutual funds at each period (SRI_alt2). Table 3.9 shows the findings are fully consistent with the baseline model (Table 3.6), with the exception of $Conv\_alt2_{LU}$, which outperforms $Conv\_alt2_{SU}$.
Third, we propose as a final exercise for the SRI category to use data on self-declared sustainable strategy. Differences between the *de jure* and *de facto* strategy in general have been well documented in the literature on mutual funds (Sensoy 2009). Recently, Candelon et al. (2018) explored the case of sustainable funds. Likewise, we use the "socially conscious" variable constructed by Morningstar from mutual funds prospectus to create an alternative SRI dummy variable SRI_alt3.

Table 3.9: Conditional effect of SRI on mutual funds performance
Alternative SRI category: Top 20% of the sustainability score

	Alpha_{t+1} I	Alpha_{t+1} II	Alpha_{t+1} III	Alpha_{t+1} IV
SRI_alt2_{SU}		-0.076**	-0.022	-0.063*
		(-2.16)	(-0.64)	(-1.82)
SRI_alt2_{LU}	0.076**		0.054*	0.013
	(2.16)		(1.91)	(0.63)
Conv_alt2_{SU}	0.022	-0.054*		-0.041*
	(0.64)	(-1.91)		(-1.69)
Conv_alt2_{LU}	0.063*	-0.013	0.041*	
	(1.82)	(-0.63)	(1.69)	
Controls	Yes	Yes	Yes	Yes
Time fixed effect	Yes	Yes	Yes	Yes
Style fixed effect	Yes	Yes	Yes	Yes
Number of Funds	2039	2039	2039	2039
Number of observations	35455	35455	35455	35455

t-statistics in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Note: Table 3.9 reports an alternative setting for Table 3.6 where the SRI funds (SRI_alt2) are in the top 20% of the sustainability score. We include time and style fixed effect and t-statistics are based on standard errors clustered at the fund level.

The results are reported in Table 3.10. Overall, the findings are less clear-cut. In particular, while SRI funds with larger universes outperform those with smaller universe, we can no more identify a difference between SRI_alt3_{SU} and Conventional mutual funds. An explanation could be that some self-declared funds use the SRI label as a marketing devise without truly enforcing it in their investment policy, making the distinction with the conventional category more blurry.

Fourth, we change our definition of investment universe. To create our category, we directly use the number of associated benchmarks instead of summing up their assets. To filter out some of the noise, we restrict our measure to benchmarks having an overlap of at least 10% with fund's portfolio. The small universe category corresponds to one or two associated benchmarks (SU_alt1). Funds with more than two associated benchmarks fall into the large universe category (LU_alt1). Table 3.11 shows that the coefficient attached to SRI_{SU_alt1} is negative and statistically significant in all regressions confirming our main conclusion.
Fifth, as discussed in section 3, we cannot rule out noise in our investment universe score.

Table 3.10: Conditional effect of SRI on mutual funds performance
Alternative SRI category: Self-declared socially conscious mutual funds

	Alpha_{t+1} I	Alpha_{t+1} II	Alpha_{t+1} III	Alpha_{t+1} IV
SRI_alt3_{SU}		-0.100**	-0.014	-0.057
		(-2.22)	(-0.36)	(-1.42)
SRI_alt3_{LU}	0.100**		0.085**	0.042
	(2.22)		(2.35)	(1.39)
Conv_alt3_{SU}	0.014	-0.085**		-0.043*
	(0.36)	(-2.35)		(-1.88)
Conv_alt3_{LU}	0.057	-0.042	0.043*	
	(1.42)	(-1.39)	(1.88)	
Controls	Yes	Yes	Yes	Yes
Time fixed effect	Yes	Yes	Yes	Yes
Style fixed effect	Yes	Yes	Yes	Yes
Number of Funds	2039	2039	2039	2039
Number of observations	35455	35455	35455	35455

t-statistics in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Note: Table 3.10 reports an alternative setting for Table 3.6 where the SRI funds (SRI_alt3) are self-declared. We include time and style fixed effect and t-statistics are based on standard errors clustered at the fund level.

Because we are not interested in the value of the score by itself, but in the resulting ranking of funds, the problem can be of limited importance. Ultimately, what we wish is to avoid misclassification (*i.e.*, include a fund in the wrong category) or to keep it as negligible as possible. In the benchmark model, mutual funds with a IU score above (below) the median fell into the large (small) universe category. Here, we use an alternative way to translate the score into the categories which should be more robust to the presence of noise. The variable (SU_alt2) takes the value of 1 a mutual fund belongs to the bottom 30% of the distribution, and 0 otherwise. The variable (LU_alt2) takes the value of 1 if the mutual fund belongs to the top 30% of the distribution, and 0 otherwise. We add for technical reasons a new category (MU_alt2) for funds located in the middle of the distribution bearing in mind that our main objective is to mitigate the risk of misclassification between the small and the large universe categories. Table 3.12 confirms our main finding as it singles out SRI funds associated with the smaller universes of investment. This category is the only one to significantly under-perform all five others at the 1% level. SRI funds having large universes perform equally to Conventional funds with a large universe.

Table 3.11: Conditional effect of SRI on mutual funds performance
Alternative IU categories: Number of associated benchmarks

	Alpha$_{t+1}$	Alpha$_{t+1}$	Alpha$_{t+1}$	Alpha$_{t+1}$
	I	II	III	IV
SRI$_{SU_alt1}$		-0.105*	-0.104**	-0.118**
		(-1.91)	(-2.00)	(-2.37)
SRI$_{LU_alt1}$	0.105*		0.001	-0.013
	(1.91)		(0.04)	(-0.41)
Conv$_{SU_alt1}$	0.104**	-0.001		-0.014
	(2.00)	(-0.04)		(-0.61)
Conv$_{LU_alt1}$	0.118**	0.013	0.014	
	(2.37)	(0.41)	(0.61)	
Controls	Yes	Yes	Yes	Yes
Time fixed effect	Yes	Yes	Yes	Yes
Style fixed effect	Yes	Yes	Yes	Yes
Number of Funds	1874	1874	1874	1874
Number of observations	31264	31264	31264	31264

t-statistics in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Note: Table 3.11 reports an alternative setting for Table 3.6 where the universe size is computed as the number of styles to which the funds as at least 10% of portfolio overlap. SU_{alt1} funds have at most 2 associated benchmarks where LU_{alt1} funds have more than 2. We include time and style fixed effect and t-statistics are based on standard errors clustered at the fund level.

Table 3.12: Conditional effect of SRI on mutual funds performance
Alternative IU categories: Extreme deciles

	Alpha_{t+1}	Alpha_{t+1}	Alpha_{t+1}	Alpha_{t+1}	Alpha_{t+1}	Alpha_{t+1}
	I	II	III	IV	V	VI
SRI_{SU_alt2}		-0.203***	-0.253***	-0.164***	-0.170***	-0.226***
		(-3.07)	(-3.73)	(-2.77)	(-2.92)	(-3.71)
SRI_{MU_alt2}	0.203***		-0.050	0.039	0.033	-0.023
	(3.07)		(-1.08)	(0.91)	(0.91)	(-0.62)
SRI_{LU_alt2}	0.253***	0.050		0.089**	0.083**	0.027
	(3.73)	(1.08)		(1.97)	(2.17)	(0.73)
Conv_{SU_alt2}	0.164***	-0.039	-0.089**		-0.006	-0.062*
	(2.77)	(-0.91)	(-1.97)		(-0.22)	(-1.85)
Conv_{MU_alt2}	0.170***	-0.033	-0.083**	0.006		-0.056**
	(2.92)	(-0.91)	(-2.17)	(0.22)		(-2.57)
Conv_{LU_alt2}	0.226***	0.023	-0.027	0.062*	0.056**	
	(3.71)	(0.62)	(-0.73)	(1.85)	(2.57)	
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Time fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Style fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Number of Funds	2039	2039	2039	2039	2039	2039
Number of observations	35455	35455	35455	35455	35455	35455

t-statistics in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Note: Table 3.12 reports an alternative setting for Table 3.6 where the universe size is 3 specific categories: SU_{alt2} (from the 1^{st} to 3^{rd} decile of the IU Score), MU_{alt2} (from the 4^{th} to 7^{th} decile of the IU Score) and LU_{alt2} (from the 8^{th} to 10^{th} decile of the IU Score). We include time and style fixed effect and t-statistics are based on standard errors clustered at the fund level.

3.6 Conclusion

We propose an empirical strategy to test the conventional wisdom stating that socially responsible investment funds (SRIs) should under-perform with respect to their conventional peers, since the former must choose securities from a smaller universe. We show that this narrative is poorly grounded in the mutual fund industry in which most mutual funds, in practice, do not consider all tradable assets when investing, but follow specific investment styles (*e.g.*, small cap or large-value) that already restrict their investment universe. Against this background, the question we ask is whether restrictions in an investment universe imposed by sustainability considerations can be compensated for in other areas such as following styles associated with a large pool of assets.

To address this question, we develop an original measure of investment universe based on a matching procedure between the portfolio holdings of US equity mutual funds and style benchmarks provided by Morningstar (*e.g.*, value, growth, large cap, financial). Morningstar benchmarks embed nearly 100% of tradable stocks in the US market. Overall, our procedure enables to recover a score, labelled *IU Score*, reflecting the size of investment universe of 2,039 US equity mutual funds, updated quarterly, from 2012Q4-2018Q4. We match this information with mutual funds sustainability score provided by Morningstar and sustainalytics to construct four categories: SRI mutual funds with smaller universe (SRI_{SU}), SRI mutual funds with larger universe (SRI_{LU}), Conventional mutual funds with smaller universe ($Conv_{SU}$) and Conventional mutual funds with larger universe($Conv_{LU}$).

Our results show that there is no difference in risk adjusted performance between SRI funds and conventional funds when we control for state-of-the-art characteristics and we consider SRI as an homogeneous group. If we break down the set of SRI mutual funds into two subgroups with respect to their universe, we show that only SRI mutual funds with the smallest universe consistently underperformed other categories. The selection of the style appears therefore as critical for SRI managers to minimize the side effect of restrictions imposed by their ethical goals. This finding is robust to a battery of robustness checks in which we alter the definition of our main categories (*i.e.*, SRI vs Conventional mutual funds and large vs small universe). We also test for reverse causality.

General Conclusion

As I have demonstrated in this thesis, active management remains a very topical subject and the debate regarding active management and managerial skills is far from being settled. The evolution of the market and notably the rise in popularity of passive strategies have forced active managers to adapt in order to remain competitive. As previously stated, this sparked a renewed interest regarding active management and the mechanisms set in place by active managers to try to outperform their peers or the market persistently. The three chapters I articulated in this thesis aimed to provide new insights on these recent development which can be broadly summarised along two main threads: (i) The ability of some funds to deviate from their peers in order to generate excess risk adjusted performance and (ii) The role of the investment universe size and strategy breadth in mitigating diversification issues. First lets recall the current state of the literature regarding these two concepts.

Peer deviation: In recent years, peer deviation (*i.e* strategy distinctiveness) has attracted its fare share of attention from academics worldwide, given that it is used as an active management proxy. Part of the literature focuses on market based data to compute such distinctiveness, either by computing the correlation of a funds returns to the average returns of its peers (see, Cremers & Petajisto 2009, Vozlyublennaia & Wu 2017) or by measuring the funds exposure to idiosyncratic risk through a multi-factor model (see, Huij & Derwall 2011, Amihud & Goyenko 2013). However, even though market based data have the benefits of being easily accessible at relatively high frequency, they cannot provide a clear picture of how managerial choices translate into distinctiveness. Hence, other studies uses portfolio holding to evaluates peer deviation (see, Kacperczyk et al. 2005, Cremers & Petajisto 2009, Choi et al. 2017). While information on holdings are much harder to obtain and to consolidate, they provide a more potent opportunity to evaluate managerial skill and strategy distinctiveness. For instance, Cremers & Petajisto (2009) propose to measure peer deviation by assessing the dissimilarity of a managers portfolio allocations with respect to the one of its referential benchmark. Similarly, Kacperczyk et al. (2005), Choi et al. (2017) propose to measure the proportion of the portfolio (*i.e.* the level of concentration) invested in specific styles , industries or countries with respect to the corresponding categories in the market portfolio. Finally, these studies provided new evidences that funds which tends to correlate less to their peers or depart from the common portfolio allocation schemes are able to outperform.

Investment universe and strategy breadth: One of the principal of modern portfolio theory pertains to the necessity to hold shares of a wide variety of individual securities to diversify away the portfolio idiosyncratic risk. Yet, in practice most managers operate on a subsection of the market on which they presumably have a better expertise: the so called “styles”. Interestingly, high profile researches in the last few years actually found evidence than managers concentrating their investment in few industries or countries where they have informational advantages are able to outperform their competitors, corroborating J.M. Keynes’s view on investment principles. These observations led new studies to investigate how some managers were able to focus their investment without worsening diversification to a point where it would become detrimental. Huij & Derwall (2011) based on Grinold & Kahn (2000) provided the first answer to the question by relating portfolio concentration to strategy breadth (*i.e* the number of separate risk exposures), they uncovered than funds concentrating their holdings exposures through multiples risk factors (namely style, sector and country risk factor) were outperforming their competitors whose exposure were either focused on one or two sources.
On a related issue, diversification or the lack thereof, has also been at the debate regarding SRI funds performance. Indeed, the restrictions imposed by ethical criteria followed by the manager could supposedly worsen portfolio’s efficiency and systematically lead ethical managers to under-perform their more conventional counterparts. Yet, contradictory evidences have been unearthed by academics worldwide, some pointing to their actual out-performance (see for instance, Nofsinger & Varma 2014), while others finding evidence of their under-performance (see, Renneboog et al. 2008*a*, El Ghoul & Karoui 2017, Ciciretti et al. 2019). Yet, the majority of the studies found no statistical evidence of difference in performance from SRI funds and non-SRI funds (Bauer et al. 2005, Barnett & Salomon 2006, Gregory & Whittaker 2007, Meziani 2014, Dolvin et al. 2019).

Each chapter developed in this thesis, though they rely on different data, models and research questions, is a variation of the two concepts described above. In what follows, I wish to re-establish where each chapter’s contribution lies and provide some avenues for future researches.
Chapter 1 proposes to take care of several empirical caveats mainly overlooked in the literature on mutual funds performance and peer distinctiveness. First, we propose to en-

dogenously determine each funds' peers using an adaptive clustering approach (AFFECT) developed by Xu et al. (2014) and then apply Sun et al. (2012)'s SDI measure to compute their strategy distinctiveness. This method relies on a limited set of assumptions and addresses issues regarding changes in the number of styles, funds' shifts in style, and the entry and exit of funds. Second, we apply Ardia & Boudt (2018) approach to formally test the difference in performance across peers, it allows to retrieve results robust to multiple testing issues and formally check whether performance is due to luck or actual skills (Barras et al. 2010). Third, to the best of our knowledge, we are the first study to assess the impact of strategy distinctiveness in the context of European EMFs. Fourth, we explore the non-linear relationship between financial performance and strategy distinctiveness and find that being too distinct or distinguishing oneself too fast from its peers (migration risk) destroy the added value of strategy distinctiveness. Turning to potential further researches, one could investigate alternatives to the SDI. Indeed the measure is simple and straightforward yet a distinctiveness measure based on portfolio average characteristics (*e.g.* TNA, manager tenure, fund's ownership , etc.) rather than returns correlation could provide new insights on the drivers of strategy distinctiveness and further consolidate the results.
Chapter 2 builds on the same premise than chapter 1, namely the intuition than peers deviation is at the heart of successful active strategies, yet deepens the analysis by focusing on the funds' portfolio holdings characteristics. More precisely, we use portfolio holdings to estimate the portfolio concentration but depart from the literature in the way we asses it. We propose to analyse portfolio concentration alongside two complementary axes : stock concentration and the number of risk exposures. While the former is straightforward to compute with a Herfindahl Index on portfolio positions (see Baks et al. 2007, Fulkerson & Riley 2019), the later requires a more careful analysis. Hence, we propose to use a novel methodology —inspired from recent developments in the risk budgeting literature— which endogenously determines uncorrelated risk factors and their relative importance to the portfolio risk exposures. The underlying idea is to distinguish truly uncorrelated sources of risk, thanks to a PCA in order to compute the strategy breadth, or in other words the diversity of the underlying strategy. The contributions are three-folds: First it provides an alternative to benchmark-centric measures, thus allowing to take into account multi-benchmarks strategies using portfolio holdings (Amihud & Goyenko 2013). Second, to the best of our knowledge this is the largest data set of European mutual fund holdings data consolidated yet, as such it allows to study a completely different set of funds and strategies than the one usually used in US focused researches. Three, directly related to the second point, the methodology is much more flexible than traditional multi-factor models (Huij & Derwall 2011, Choi et al. 2017) to accommodate a vast diversity of funds investing either globally, internationally

or domestically, as it is the case in the EU market. Or main results highlights that funds concentrating their holdings yet spreading their risk exposures on multiple uncorrelated risk sources are significantly outperforming the rest of the market, thus alleviating the diversification issue linked to a concentrated portfolio. However, one remaining drawback pertains to the identification of the uncorrelated endogenous risk factors. Further studies could provide an alternative to the PCA to uncover the endogenous factors and link them to worldwide economic-based ones. We suggest looking at Meucci et al. (2015) latest work on minimum torsion bets which provide an interesting framework to do so.
Finally Chapter 3 investigates the idea than SRI funds are subject to under-performance relatively to common funds as their investment universe is restricted. However, as stated above most funds already operates on sub-market segments, this raising the question: why should an ethical restriction be any different than a style restriction? To provide an answer to this question, we propose a novel methodology to assess the investment universe size and label our measure the IU Score. Our hypothesis is that —similarly to other investment constraints— ethical restrictions might reduce one's investment universe, yet that some managers are able to mitigate this issue by investing in larger style (providing a sufficient amount of stock opportunities) or across multiple ones. Similarly to chapter 2, we use portfolio holdings information to assess how managers may alleviate concentration issues (here restricted universes). However, in this study we propose to compute the investment universe size, an observe whether it allows to maintain a sufficiently high level of diversification for the funds to remain competitive. Our results highlight that SRI funds having a small investment universe are indeed under-performing the rest of the market, while their counterpart with larger universes are performing as well as conventional funds with large universes and even outperforming conventional funds with small investment universes. As such, our results allow to link seemingly contradictory results found in the literature, by introducing the investment universe size. Given that the literature on ethical funds is much more recent, there are still many research avenues to investigate. First, expanding the analysis the non-US funds could provide more information on whether or not the current academic consensus on SRI performance is driven by US specific characteristics. Second, the liquidity dimension of such funds has been overlooked in the literature and could provide very interesting new insights on SRI vs Conventional funds characteristics. Third, the difference in exposition to quality factors may also be further examined to explain SRI funds performance, as their ethical criteria are supposedly driving them to high quality firms.

To summarise, the present thesis not only provides new evidence of skilful active management, but provides alternatives tools to distinguish it. On the one hand, we argue that it allows academics to build on these novel measures to unearth new market dynamics, and on the other hand investors to distinguish different drivers of successful active strategies more effectively.

Bibliography

Almazan, A., Brown, K. C., Carlson, M. & Chapman, D. A. (2004), 'Why constrain your mutual fund manager?', *Journal of Financial Economics* **73**(2), 289–321.

Amihud, Y. & Goyenko, R. (2013), 'Mutual fund's r 2 as predictor of performance', *The Review of Financial Studies* **26**(3), 667–694.

Ammann, M., Bauer, C., Fischer, S. & Müller, P. (2019), 'The impact of the morningstar sustainability rating on mutual fund flows', *European Financial Management* **25**(3), 520–553.

Ardia, D. & Boudt, K. (2018), 'The peer performance ratios of hedge funds', *Journal of Banking & Finance* **87**, 351–368.

Auer, B. R. (2016), 'Do socially responsible investment policies add or destroy european stock portfolio value?', *Journal of business ethics* **135**(2), 381–397.

Baks, K., Busse, J. & Green, C. (2007), 'Fund managers who take big bets: Skilled or overconfident', *Paper SSRN* .

Banegas, A., Gillen, B., Timmermann, A. & Wermers, R. (2013), 'The cross section of conditional mutual fund performance in european stock markets', *Journal of Financial economics* **108**(3), 699–726.

Banerjee, A., Dhillon, I., Ghosh, J., Merugu, S. & Modha, D. S. (2007), 'A generalized maximum entropy approach to bregman co-clustering and matrix approximation', *Journal of Machine Learning Research* **8**(Aug), 1919–1986.

Banz, R. W. (1981), 'The relationship between return and market value of common stocks', *Journal of financial economics* **9**(1), 3–18.

Barberis, N., Shleifer, A. & Wurgler, J. (2005), 'Comovement', *Journal of financial economics* **75**(2), 283–317.

Barnett, M. L. & Salomon, R. M. (2006), 'Beyond dichotomy: The curvilinear relationship between social responsibility and financial performance', *Strategic management journal* **27**(11), 1101–1122.

Barras, L., Scaillet, O. & Wermers, R. (2010), 'False discoveries in mutual fund performance: Measuring luck in estimated alphas', *The journal of finance* **65**(1), 179–216.

Basu, S. (1977), 'Investment performance of common stocks in relation to their price-earnings ratios: A test of the efficient market hypothesis', *The journal of Finance* **32**(3), 663–682.

Bauer, R., Derwall, J. & Otten, R. (2007), 'The ethical mutual fund performance debate: New evidence from canada', *Journal of Business Ethics* **70**(2), 111–124.

Bauer, R., Koedijk, K. & Otten, R. (2005), 'International evidence on ethical mutual fund performance and investment style', *Journal of Banking & Finance* **29**(7), 1751–1767.

BCG (2010), 'In search of stable growth: Global asset management 2010'.
URL: *https://www.bcg.com/documents/file53448.pdf*

BCG (2019), 'Global asset management 2019: Will these '20s roar?'.
URL: *https://www.bcg.com/publications/2019/global-asset-management-will-these-20s-roar.aspx*

Ben-David, I., Li, J., Rossi, A. & Song, Y. (2019), 'What do mutual fund investors really care about?', *Fisher College of Business Working Paper* (2019-03), 005.

Béreau, S., Casteleyn, A., Gnabo, J.-Y. & Zwinkels, R. (2015), Commonality among equity mutual funds: impact on financial stability and price discovery, *in* 'Working Paper UNamur'.

Berk, J. B. & Green, R. C. (2004), 'Mutual fund flows and performance in rational markets', *Journal of political economy* **112**(6), 1269–1295.

Borgers, A., Derwall, J., Koedijk, K. & ter Horst, J. (2015), 'Do social factors influence investment behavior and performance? evidence from mutual fund holdings', *Journal of Banking & Finance* **60**, 112–126.

Boyle, P., Garlappi, L., Uppal, R. & Wang, T. (2012), 'Keynes meets markowitz: The trade-off between familiarity and diversification', *Management Science* **58**(2), 253–272.

Brands, S., Brown, S. J. & Gallagher, D. R. (2005), 'Portfolio concentration and investment manager performance', *International Review of Finance* **5**(3-4), 149–174.

Brown, S. J. & Goetzmann, W. N. (1997), 'Mutual fund styles', *Journal of financial Economics* **43**(3), 373–399.

Brown, S. J. & Goetzmann, W. N. (2001), Hedge funds with style, Technical report, National Bureau of Economic Research.

Bussière, M., Hoerova, M. & Klaus, B. (2015), 'Commonality in hedge fund returns: Driving factors and implications', *Journal of Banking & Finance* **54**, 266–280.

Candelon, B., Hasse, J.-B. & Lajaunie, Q. (2018), Sri: Truths and lies, Technical report, Université catholique de Louvain, Center for Operations Research and Econometrics.

Carhart, M. M. (1997), 'On persistence in mutual fund performance', *The Journal of finance* **52**(1), 57–82.

Ceccarelli, M., Ramelli, S. & Wagner, A. F. (2019), 'When investors call for climate responsibility, how do mutual funds respond?', *Swiss Finance Institute Research Paper* (19-13).

Chen, J., Hong, H., Huang, M. & Kubik, J. D. (2004), 'Does fund size erode mutual fund performance? the role of liquidity and organization', *American Economic Review* **94**(5), 1276–1302.

Choi, N., Fedenia, M., Skiba, H. & Sokolyk, T. (2017), 'Portfolio concentration and performance of institutional investors worldwide', *Journal of Financial Economics* **123**(1), 189–208.

Ciciretti, R., Dalo, A. & Dam, L. (2019), 'The contributions of betas versus characteristics to the esg premium'.

Clark, G. L., Feiner, A. & Viehs, M. (2015), 'From the stockholder to the stakeholder: How sustainability can drive financial outperformance', *Available at SSRN 2508281* .

Clark, T. E. & McCracken, M. W. (2009), 'Improving forecast accuracy by combining recursive and rolling forecasts', *International Economic Review* **50**(2), 363–395.

Clarke, R., De Silva, H. & Thorley, S. (2002), 'Portfolio constraints and the fundamental law of active management', *Financial Analysts Journal* **58**(5), 48–66.

Cohen, R. B., Coval, J. D. & Pástor, L. (2005), 'Judging fund managers by the company they keep', *The Journal of Finance* **60**(3), 1057–1096.

Cremers, K. M., Fulkerson, J. A. & Riley, T. B. (2019), 'Challenging the conventional wisdom on active management: A review of the past 20 years of academic literature on actively managed mutual funds', *Financial Analysts Journal* **75**(4), 8–35.

Cremers, K. M. & Petajisto, A. (2009), 'How active is your fund manager? a new measure that predicts performance', *The Review of Financial Studies* **22**(9), 3329–3365.

Daniel, K., Grinblatt, M., Titman, S. & Wermers, R. (1997*a*), 'Measuring mutual fund performance with characteristic-based benchmarks', *The Journal of finance* **52**(3), 1035–1058.

Daniel, K., Grinblatt, M., Titman, S. & Wermers, R. (1997*b*), 'Measuring mutual fund performance with characteristic-based benchmarks', *The Journal of finance* **52**(3), 1035–1058.

Del Guercio, D. & Tkac, P. A. (2008), 'Star power: The effect of monrningstar ratings on mutual fund flow', *Journal of Financial and Quantitative Analysis* **43**(4), 907–936.

Deli, D. N. & Varma, R. (2002), 'Contracting in the investment management industry:: evidence from mutual funds', *Journal of Financial Economics* **63**(1), 79–98.

Deloitte (2019), '2019 banking and capital markets outlook: Reimagining transformation'.
URL: *https://www2.deloitte.com/content/dam/Deloitte/global/Documents/Financial-Services/gx-fsi-dcfs-2019-banking-cap-markets-outlook.pdf*

Derwall, J., Koedijk, K. & Ter Horst, J. (2011), 'A tale of values-driven and profit-seeking social investors', *Journal of Banking & Finance* **35**(8), 2137–2147.

DiBartolomeo, D. & Witkowski, E. (1997), 'Mutual fund misclassification: Evidence based on style analysis', *Financial Analysts Journal* **53**(5), 32–43.

Dolvin, S., Fulkerson, J. & Krukover, A. (2019), 'Do "good guys" finish last? the relationship between morningstar sustainability ratings and mutual fund performance', *The Journal of Investing* **28**(2), 77–91.

EFAMA (2017), 'Asset management in europe: An overview of the asset management industry with a special section on the capital markets union'.
URL: *https://www.efama.org*

El Ghoul, S. & Karoui, A. (2017), 'Does corporate social responsibility affect mutual fund performance and flows?', *Journal of Banking & Finance* **77**, 53–63.

Elton, E. J., Gruber, M. J. & Blake, C. R. (2013), 'The performance of separate accounts and collective investment trusts', *Review of Finance* **18**(5), 1717–1742.

EY (2017), 'Reshaping around the investor: Global etf research 2017'.
URL: *https://www.ey.com/Publication/vwLUAssets/ey-global-etf-survey-2017$FILE/ey-global-etf-survey-2017.pdf*

EY (2020), 'Understanding luxembourg fund in 10 minutes'.
URL: *https://www.ey.com/Publication/vwLUAssets/ey-understanding-luxembourg-fund-in-10-minutes-en/$File/ey-understanding-luxembourg-fund-in-10-minutes-en.pdf*

Fama, E. F. & French, K. R. (1993), 'Common risk factors in the returns on stocks and bonds', *Journal of* .

Fama, E. F. & French, K. R. (1998), 'Value versus growth: The international evidence', *The journal of finance* **53**(6), 1975–1999.

Fama, E. F. & French, K. R. (2010), 'Luck versus skill in the cross-section of mutual fund returns', *The journal of finance* **65**(5), 1915–1947.

Fama, E. F. & French, K. R. (2017), 'International tests of a five-factor asset pricing model', *Journal of financial Economics* **123**(3), 441–463.

Ferreira, M. A., Keswani, A., Miguel, A. F. & Ramos, S. B. (2013), 'The determinants of mutual fund performance: A cross-country study', *Review of Finance* **17**(2), 483–525.

Ferreira, M. A., Matos, P. & Pires, P. (2018), 'Asset management within commercial banking groups: International evidence', *The Journal of Finance* **73**(5), 2181–2227.

Franck, A. & Kerl, A. (2013), 'Analyst forecasts and european mutual fund trading', *Journal of Banking & Finance* **37**(8), 2677–2692.

FSB (2019), 'Global monitoring report on non-bank financial intermediation 2019'.
URL: *https://www.fsb.org/wp-content/uploads/P190120.pdf*

Fulkerson, J. A. & Riley, T. B. (2019), 'Portfolio concentration and mutual fund performance', *Journal of Empirical Finance* **51**, 1–16.

Graef, F., Vogt, P., Vonhoff, V. & Weigert, F. (2018), 'Cash holdings and the performance of european mutual funds'.

Graef, F., Vogt, P., Vonhoff, V. & Weigert, F. (2019), 'Cash holdings and the performance of european mutual funds', *Finance Research Letters* **29**, 285–291.

Gregory, A. & Whittaker, J. (2007), 'Performance and performance persistence of "ethical" unit trusts in the uk', *Journal of Business Finance & Accounting* **34**(7-8), 1327–1344.

Grinold, R. C. (1989), 'The fundamental law of active management', *The Journal of Portfolio Management* **15**(3), 30–37.

Grinold, R. C. & Kahn, R. N. (2000), 'Active portfolio management'.

Gruber, M. J. (1996), 'Another puzzle: The growth in actively managed mutual funds', *The journal of finance* **51**(3), 783–810.

Gupta-Mukherjee, S. (2013), 'When active fund managers deviate from their peers: Implications for fund performance', *Journal of Banking & Finance* **37**(4), 1286–1305.

Hartigan, J. A. (1975), *Clustering algorithms*, Vol. 209, Wiley New York.

Hartzmark, S. M. & Sussman, A. B. (2019), 'Do investors value sustainability? a natural experiment examining ranking and fund flows', *The Journal of Finance* **74**(6), 2789–2837.

Haslem, J. (2009), *Mutual funds: risk and performance analysis for decision making*, John Wiley & Sons.

Haslem, J. A. (2017), 'Mutual funds: Risk and performance analysis for decision making', *Available at SSRN 2026563* .

Hoberg, G., Kumar, N. & Prabhala, N. (2017), 'Mutual fund competition, managerial skill, and alpha persistence', *The Review of Financial Studies* **31**(5), 1896–1929.

Hoepner, A. G. F. (2010), 'Portfolio diversification and environmental, social or governance criteria: Must responsible investments really be poorly diversified?', *Working paper, University of Reading, United Nations Principles for Responsible Investment* .

Huebscher, R. (June 2019), 'The disappointing reality about esg fund performance'.
URL: *https://www.advisorperspectives.com/articles/2019/06/24/the-disappointing-reality-about-esg-fund-performance*

Huij, J. & Derwall, J. (2011), 'Global equity fund performance, portfolio concentration, and the fundamental law of active management', *Journal of Banking & Finance* **35**(1), 155–165.

Hunter, D., Kandel, E., Kandel, S. & Wermers, R. (2014), 'Mutual fund performance evaluation with active peer benchmarks', *Journal of Financial economics* **112**(1), 1–29.

ICI (2017), 'Investment company institute - release: Worldwide regulated open-end fund assets and flows, first quarter 2017'.
URL: *https://www.ici.org/pdf/2017$_f$actbook.pdf*

Jegadeesh, N. & Titman, S. (1993), 'Returns to buying winners and selling losers: Implications for stock market efficiency', *The Journal of finance* **48**(1), 65–91.

Jensen, M. C. (1968), 'The performance of mutual funds in the period 1945–1964', *The Journal of finance* **23**(2), 389–416.

Jensen, M. C. (1969), 'Risk, the pricing of capital assets, and the evaluation of investment portfolios', *Journal of business* **42**(2), 167–247.

Joliet, R. & Titova, Y. (2018), 'Equity sri funds vacillate between ethics and money: An analysis of the funds' stock holding decisions', *Journal of Banking & Finance* **97**, 70–86.

Jolliffe, I. T. & Cadima, J. (2016), 'Principal component analysis: a review and recent developments', *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **374**(2065), 20150202.

Kacperczyk, M., Nieuwerburgh, S. V. & Veldkamp, L. (2014), 'Time-varying fund manager skill', *The Journal of Finance* **69**(4), 1455–1484.

Kacperczyk, M., Sialm, C. & Zheng, L. (2005), 'On the industry concentration of actively managed equity mutual funds', *The Journal of Finance* **60**(4), 1983–2011.

Kashima, H., Hu, J., Ray, B. & Singh, M. (2008), K-means clustering of proportional data using l1 distance, *in* 'Pattern Recognition, 2008. ICPR 2008. 19th International Conference on', IEEE, pp. 1–4.

Keynes, J. M., Johnson, E. & Moggridge, D. (1983), 'Keynes as an investor', *The Collected Works of John Maynard Keynes* **12**, 1–113.

Koski, J. L. & Pontiff, J. (1999), 'How are derivatives used? evidence from the mutual fund industry', *The journal of finance* **54**(2), 791–816.

Lang, G. & Köhler, M. (2011), How does the domiciliation decision affect mutual fund fees?, Technical report, ZEW Discussion Papers.

Lessard, D. R. (1974), 'World, national, and industry factors in equity returns', *The Journal of Finance* **29**(2), 379–391.

Levy, A. & Livingston, M. (1995), *The gains from diversification reconsidered: Transaction costs and superior information*, Blackwell Publishers.

Lintner, J. (1965), 'Security prices, risk, and maximal gains from diversification', *The journal of finance* **20**(4), 587–615.

Loeb, G. M. (2007), *Battle for investment survival*, Vol. 36, John Wiley & Sons.

Malkiel, B. G. (1995), 'Returns from investing in equity mutual funds 1971 to 1991', *The Journal of finance* **50**(2), 549–572.

Mao, J. & Jain, A. K. (1996), 'A self-organizing network for hyperellipsoidal clustering (hec)', *Ieee transactions on neural networks* **7**(1), 16–29.

Markowitz, H. (1952), 'Portfolio selection', *The journal of finance* **7**(1), 77–91.

Matias, C. & Miele, V. (2017), 'Statistical clustering of temporal networks through a dynamic stochastic block model', *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **79**(4), 1119–1141.

Merton, R. C. (1987), 'A simple model of capital market equilibrium with incomplete information', *The journal of finance* **42**(3), 483–510.

Meucci, A. (2010), 'Managing diversification', *Risk* **22**(5), 74.

Meucci, A., Santangelo, A. & Deguest, R. (2015), 'Risk budgeting and diversification based on optimised uncorrelated factors', *Risk* **11**(1), 70–75.

Meziani, A. S. (2014), 'Investing with environmental, social, and governance issuesin mind: From the back to the fore of style investing', *The Journal of Investing* **23**(3), 115–124.

Morningstar (2011), 'A brief history of indexing'.
URL: *https://www.morningstar.com/articles/390749/a-brief-history-of-indexing*

Morningstar (2016), 'Global investors shun equity, return to fixed income in 2016 2016 global asset flows report'.
URL: *https://www.vfb.be/vfb/Media/Default/news/Morningstar.pdf*

Mossin, J. (1966), 'Equilibrium in a capital asset market', *Econometrica: Journal of the econometric society* pp. 768–783.

Nofsinger, J. & Varma, A. (2014), 'Socially responsible funds and market crises', *Journal of Banking & Finance* **48**, 180–193.

Otten, R. & Bams, D. (2002), 'European mutual fund performance', *European Financial Management* **8**(1), 75–101.

Papke, L. E. & Wooldridge, J. M. (1996), 'Econometric methods for fractional response variables with an application to 401 (k) plan participation rates', *Journal of applied econometrics* **11**(6), 619–632.

Patel, S. & Sarkissian, S. (2017), 'To group or not to group? evidence from mutual fund databases', *Journal of Financial and Quantitative Analysis* **52**(5), 1989–2021.

Porter, G. E. & Trifts, J. W. (2014), 'The career paths of mutual fund managers: the role of merit', *Financial Analysts Journal* **70**(4), 55–71.

Pozen, R. C. & Crane, S. D. (1998), *The mutual fund business*, the MIT Press.

PWC (2018), 'Benchmark your global fund distribution 2018'.
URL: *https://www.pwc.lu/en/fund-distribution/docs/pwc-publ-gfd-march-2018.pdf*

Reed, W. R. (2015), 'On the practice of lagging variables to avoid simultaneity', *Oxford Bulletin of Economics and Statistics* **72**(6), 897–905.

Renneboog, L., Ter Horst, J. & Zhang, C. (2008*a*), 'The price of ethics and stakeholder governance: The performance of socially responsible mutual funds', *Journal of corporate finance* **14**(3), 302–322.

Renneboog, L., Ter Horst, J. & Zhang, C. (2008*b*), 'Socially responsible investments: Institutional aspects, performance, and investor behavior', *Journal of banking & finance* **32**(9), 1723–1742.

Richey, G. M. (2016), 'Sin is in: An alternative to socially responsible investing?', *The Journal of Investing* **25**(2), 136–143.

Riedl, A. & Smeets, P. (2017), 'Why do investors hold socially responsible mutual funds?', *The Journal of Finance* **72**(6), 2505–2550.

Roncalli, T. (2013), *Introduction to risk parity and budgeting*, CRC Press.

Roncalli, T. & Weisang, G. (2016), 'Risk parity portfolios with risk factors', *Quantitative Finance* **16**(3), 377–388.

Rosen, B. N., Sandler, D. M. & Shani, D. (1991), 'Social issues and socially responsible investment behavior: A preliminary empirical investigation', *Journal of Consumer Affairs* **25**(2), 221–234.

Sensoy, B. A. (2009), 'Performance evaluation and self-designated benchmark indexes in the mutual fund industry', *Journal of Financial Economics* **92**(1), 25–39.

Sharpe, W. F. (1964), 'Capital asset prices: A theory of market equilibrium under conditions of risk', *The journal of finance* **19**(3), 425–442.

Sharpe, W. F. (1991), 'The arithmetic of active management', *Financial Analysts Journal* **47**(1), 7–9.

Shukla, R. K. & Van Inwegen, G. B. (1995), 'Do locals perform better than foreigners?: An analysis of uk and us mutual fund managers', *Journal of Economics and Business* **47**(3), 241–254.

Statman, M. (1987), 'How many stocks make a diversified portfolio?', *Journal of financial and quantitative analysis* **22**(3), 353–363.

Statman, M. (2000), 'Socially responsible mutual funds (corrected)', *Financial Analysts Journal* **56**(3), 30–39.

Statman, M. (2004), 'The diversification puzzle', *Financial Analysts Journal* **60**(4), 44–53.

Stein, J. C. (2002), 'Information production and capital allocation: Decentralized versus hierarchical firms', *The journal of finance* **57**(5), 1891–1921.

Sun, Z., Wang, A. & Zheng, L. (2012), 'The road less traveled: Strategy distinctiveness and hedge fund performance', *Review of Financial Studies* **25**(1), 96–143.

Treynor, J. L. (1961), *Toward a theory of market value of risky assets*.

Trinks, P. J. & Scholtens, B. (2017), 'The opportunity cost of negative screening in socially responsible investing', *Journal of Business Ethics* **140**(2), 193–208.

Tsai, H.-J. & Wu, Y. (2015), 'Performance of foreign and global mutual funds: The role of security selection, region-shifting, and style-shifting abilities', *Financial Review* **50**(4), 517–545.

Van Nieuwerburgh, S. & Veldkamp, L. (2009), 'Information immobility and the home bias puzzle', *The Journal of Finance* **64**(3), 1187–1215.

Verbeek, M. (2008), *A guide to modern econometrics*, John Wiley & Sons.

Vozlyublennaia, N. & Wu, Y. (2017), 'Mutual funds apart from the crowd', *Paper SSRN* .

Wermers, R. (2000), 'Mutual fund performance: An empirical decomposition into stock-picking talent, style, transactions costs, and expenses', *The Journal of Finance* **55**(4), 1655–1695.

Whelan, K. (2014), 'Ireland's economic crisis: The good, the bad and the ugly', *Journal of Macroeconomics* **39**, 424–440.

Xu, K. S., Kliger, M. & Hero Iii, A. O. (2014), 'Adaptive evolutionary clustering', *Data Mining and Knowledge Discovery* **28**(2), 304–336.

Yang, Y., Harris, D. P., Luo, F., Xiong, W., Joachimiak, M., Wu, L., Dehal, P., Jacobsen, J., Yang, Z., Palumbo, A. V. et al. (2009), 'Snapshot of iron response in shewanella oneidensis by gene network reconstruction', *BMC genomics* **10**(1), 131.

Zhu, Q. (2020), 'The missing new funds', *Management Science* .

Zivot, E. & Wang, J. (2006), 'Modelling financial time series with s-plus', *Springer* pp. 429–478.

Appendices

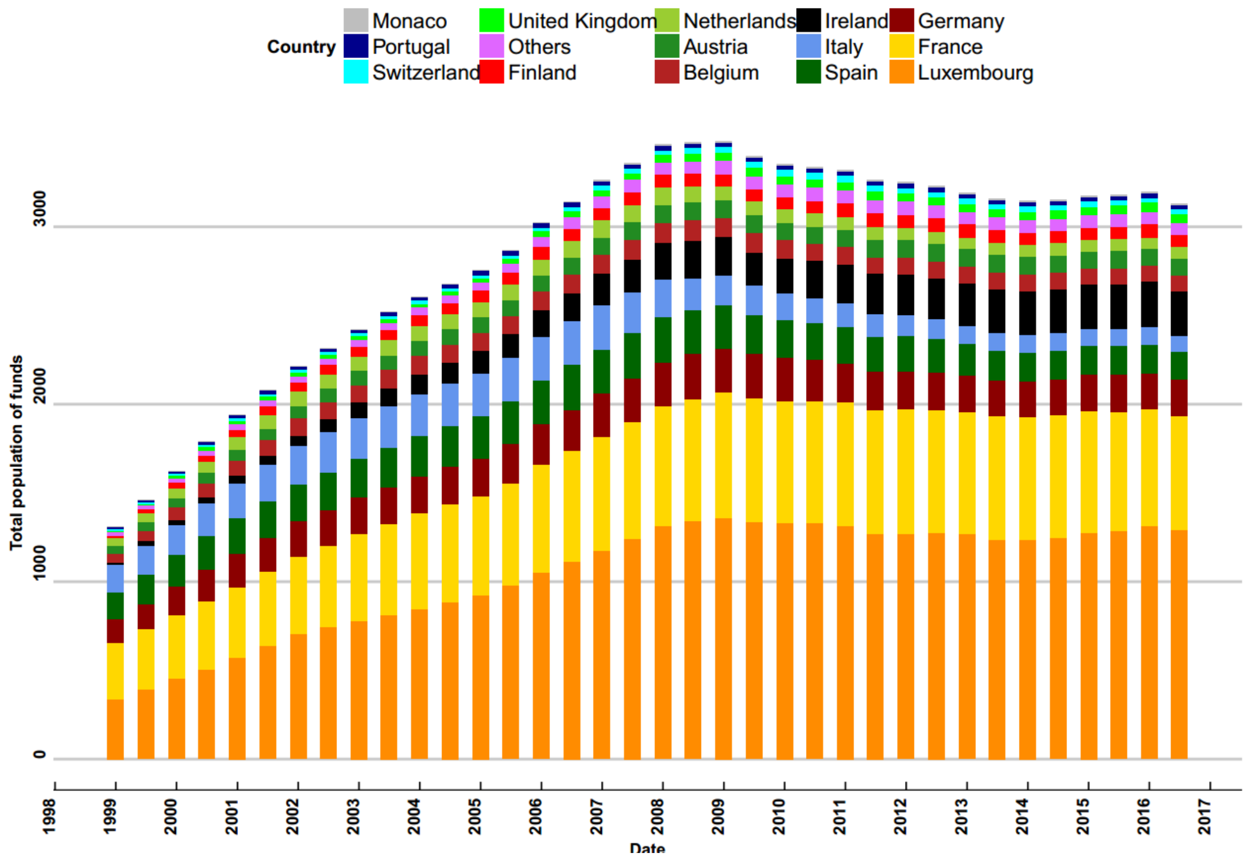
Appendix A: Overview of the EEMF market

A.1: Country-based analysis

Figure 1.1 shows the number of active mutual funds comprising our sample over time and dispatched by countries. A visual inspection shows marked differences in our sample, with Luxembourg and France accounting for more than 50% of EEMFs (see also Table 1.3). On the contrary, Portugal, Sweden, and Slovenia together account for only 2%. These percentages are in line with Graef et al. (2018) based on 1,464 funds, and we account for 4,957. The equity mutual funds' domicile used in the study is based on Morningstar.

With this in mind, we turn to the evolution of the EEMF market and note both the similarities and the differences across countries. First, taken as a whole, the industry displays two separate phases. The first is an expansion phase from 1999 to the end of 2008. During this period, the number of funds in Europe increased sharply to reach an all-time high of 3,497 from 1,310. This trend subsequently vanishes, and the industry enters a stabilizing or slightly contracting phase. In December 2016, there were 3,049 funds, down by 13% from the historical peak. A more careful inspection enables us to separate contrasting behaviors across countries. The global picture is partially driven by Luxembourg, in which the expansion phase was particularly marked. The same applies to Ireland. Both countries experienced strong growth in the number of local funds that materialized into significant market share gains. Figure 1.2 displays the market share per country over time. France and Germany maintained a fairly similar market share throughout our sample. No comparable pattern is observable for Italy and Spain, which lost some ground relative to other nations during 1999–2016. Following 2009, the number of funds decreased in Italy and Spain, remained fairly stable in Luxembourg, Germany, and France, and increased slightly in Ireland.

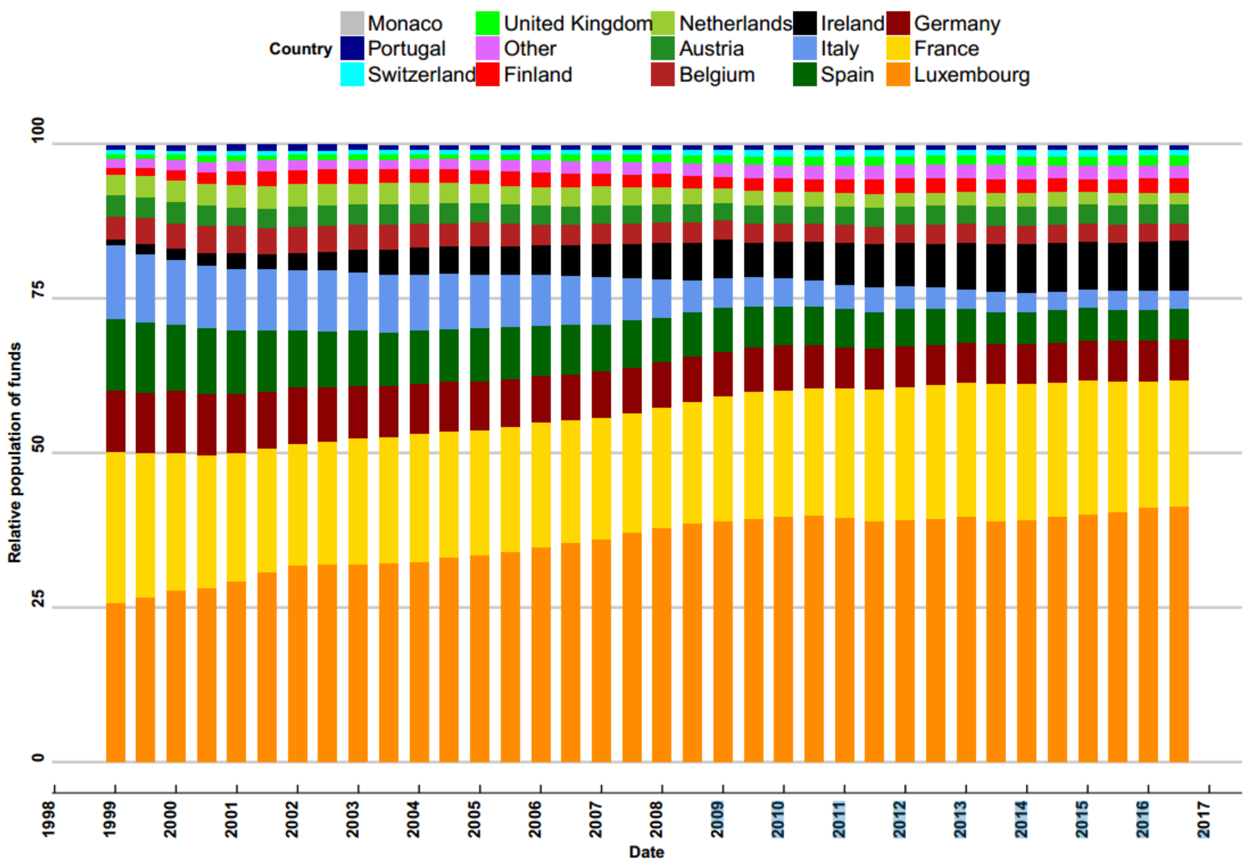
Figure 1.1: Total population of funds by country



Note: Figure 1.1 reports the total number of funds by country from 1999 to 2016.

Before proceeding with our analysis, it is worth noting that we cannot, using our data, formally identify the underlying factors driving the diverging trends. Nevertheless, from the literature on mutual fund developments, we can recall some established features and overcome them with respect to our figures. Traditionally, the mutual funds industry has responded to various factors (e.g., capital inflows from investors) that can be sensitive to competition from alternative investment vehicles such as Exchange traded funds and the market environment. The two factors are drivers of the liability and asset sides of mutual funds (Vozlyublennaia & Wu 2017). More recently, changes in regulations have also been mentioned as determinants of mutual fund development amid the implementation of the MiFiD II regulation that has been blamed by the industry for dramatically raising operating costs and penalizing small-scale structures. Our data show that the dates of the successive turning points correspond to financial events consistent with some of the drivers previously mentioned, especially the burst that caused the so-called European debt crisis. Moreover, the 2007–2009 financial crisis rooted in the U.S. economy had a significant impact on the

Figure 1.2: Relative population of funds by country



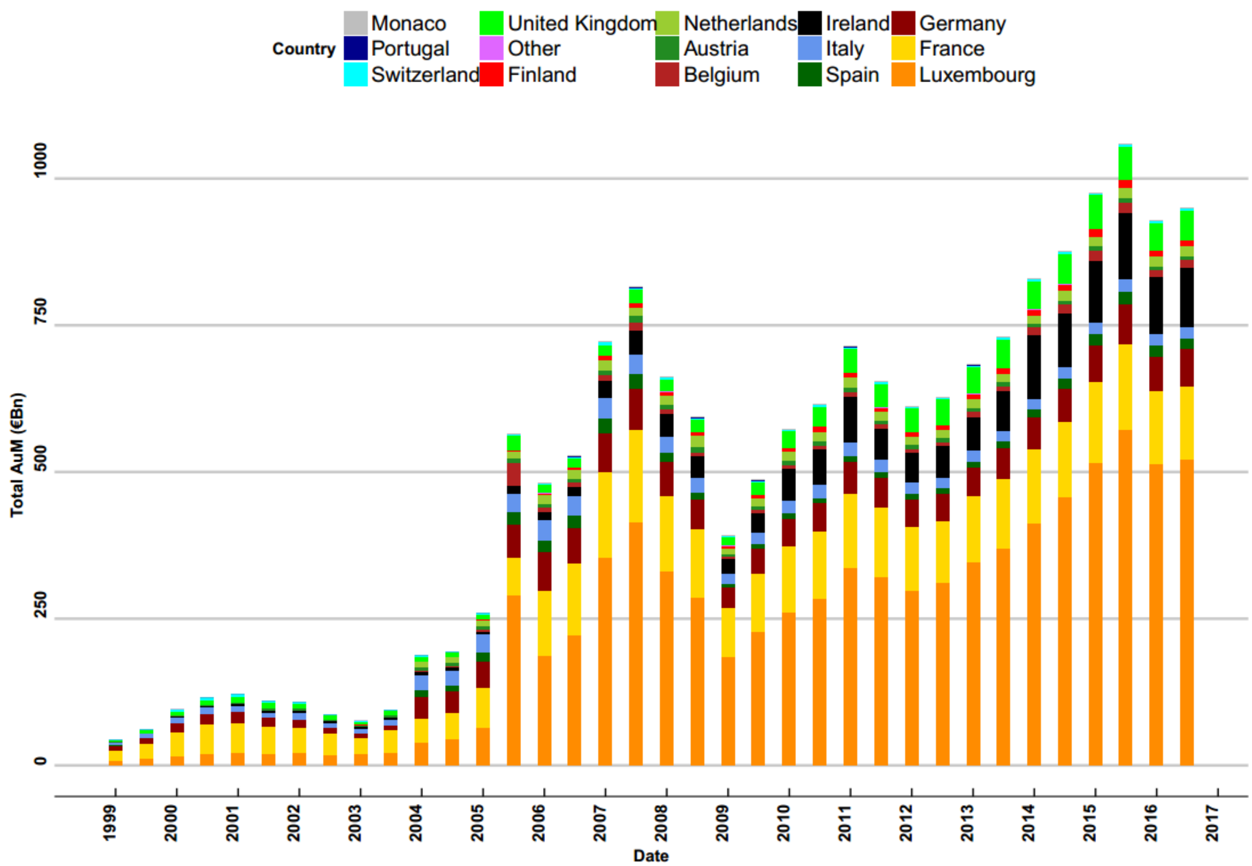
Note: Figure 1.2 reports the share of funds by country from 1999 to 2016.

EEMF market when considering the number of active mutual funds. In 2016, there was a slight decrease in the number of mutual funds that could correspond to the implementation of regulatory reforms within MiFiD. However, in this descriptive phase, we cannot formally isolate the effect of regulation from other factors. Regarding cross-country comparisons, one way to interpret the heterogeneity in our data is to consider the location of mutual fund investments. Our sample includes both mutual funds engaged primarily in foreign investments and funds with an informational advantage engaged in local investments[29]. For instance, mutual funds in Luxembourg are known to follow a worldwide strategy (Lang & Köhler 2011). By contrast, Italy and Spain host mutual funds investing both abroad and locally. As a result, the strong impact of the European crisis on these two countries may have spilled over to local mutual funds. Mutual funds in Luxembourg, along with France and Germany, resisted the European turbulence more forcefully because of either lower dependence on the local market or the better resilience of their domestic economy to the crisis. Interestingly, Ireland, whose economy was dramatically hurt in 2008 and 2010 (Whelan 2014), experienced an expansion of mutual fund activity. This finding supports the notion that mutual funds with an international investment strategy were better protected between 2008 and 2011.

Figures 1.3 and 1.4 complete the picture by considering total AUM as well as the average size of equity mutual funds per country. We now analyze whether a particular segment of the market and, specifically, small-scale funds, were more affected by successive changes in the economic and regulatory environment. As expected, the evolution of total AUM is more volatile than the number of funds. Although remaining relatively low, the AUM increased dramatically in 2005 and continued to grow until the end of 2007. The occurrence of the U.S. financial crisis seems to have reverted the trend because AUM sharply declined from approximately €800 billion in 2007 to approximately €275 billion in 2009. Next, the EEMF industry resumed its growth until 2016 with two temporary episodes of slight decreases in 2011 and 2016. Figure 4 reports the average fund size. We observe a marked difference from the previous graph. In particular, from 1999 to 2002, although the industry was smaller than the rest of the sample, average size was large. Therefore, this period was marked by a concentration of the market with a limited number of large players. Unsurprisingly, the crisis affected the stock market and value of the EEMF portfolio. Average size rebounded early during the crisis and started to grow again from 2009. Regarding total AUM, the upward trend was halted twice, in 2011 and in 2016. Overall, no evidence from these results show that the architecture of the mutual funds industry has been markedly distorted across

[29] As discussed in the literature, mutual funds possess local knowledge and privileged contacts with companies and market participants (see Shukla & Van Inwegen (1995)).

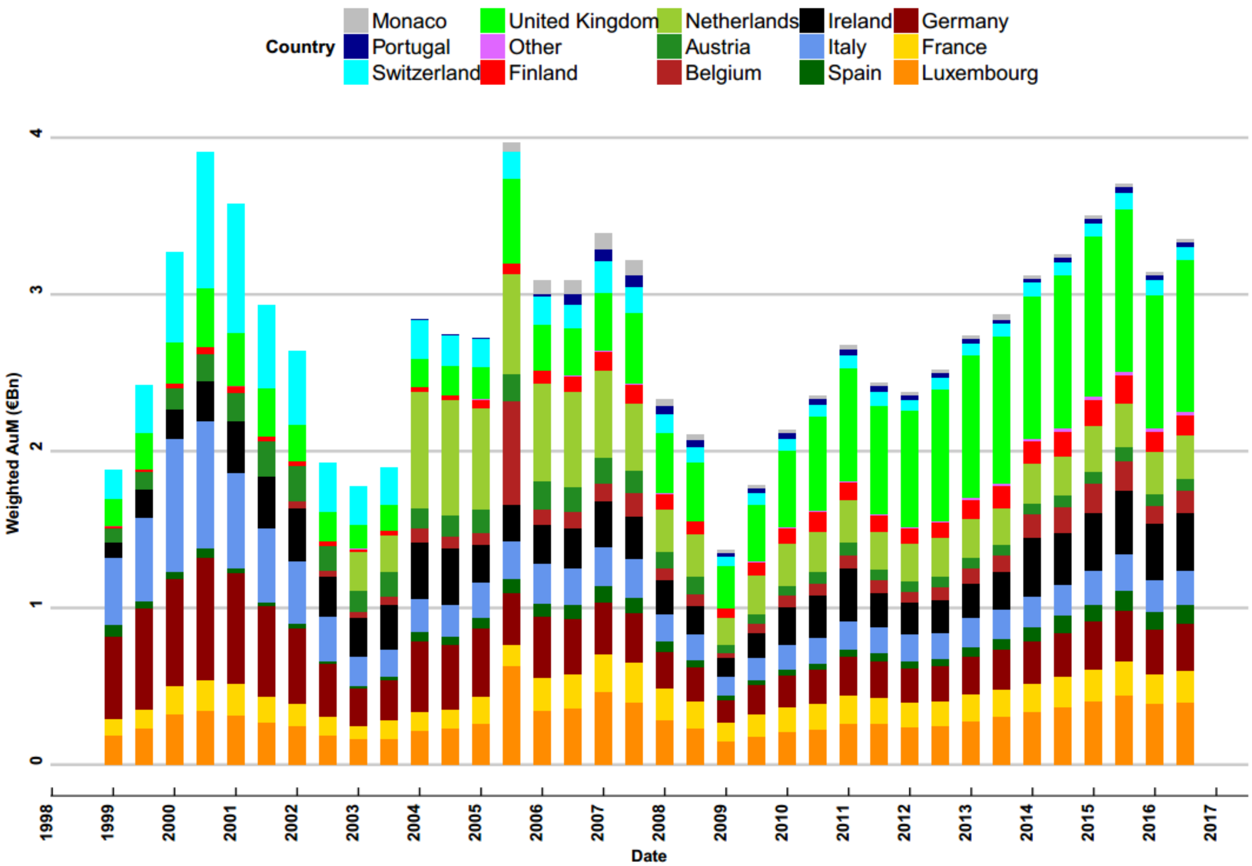
Figure 1.3: Total TNA per country



Note: Figure 1.3 reports the total TNA (in €Bn) of funds by country from 1999 to 2016.

time. There is widespread concern in the industry that increasing demand from regulators might lead to the consolidation of the global funds industry to maintain profitability. In our data, we do not pick up on an eviction of small funds following the implementation of new regulations.

Figure 1.4: Relative TNA per country's population of funds

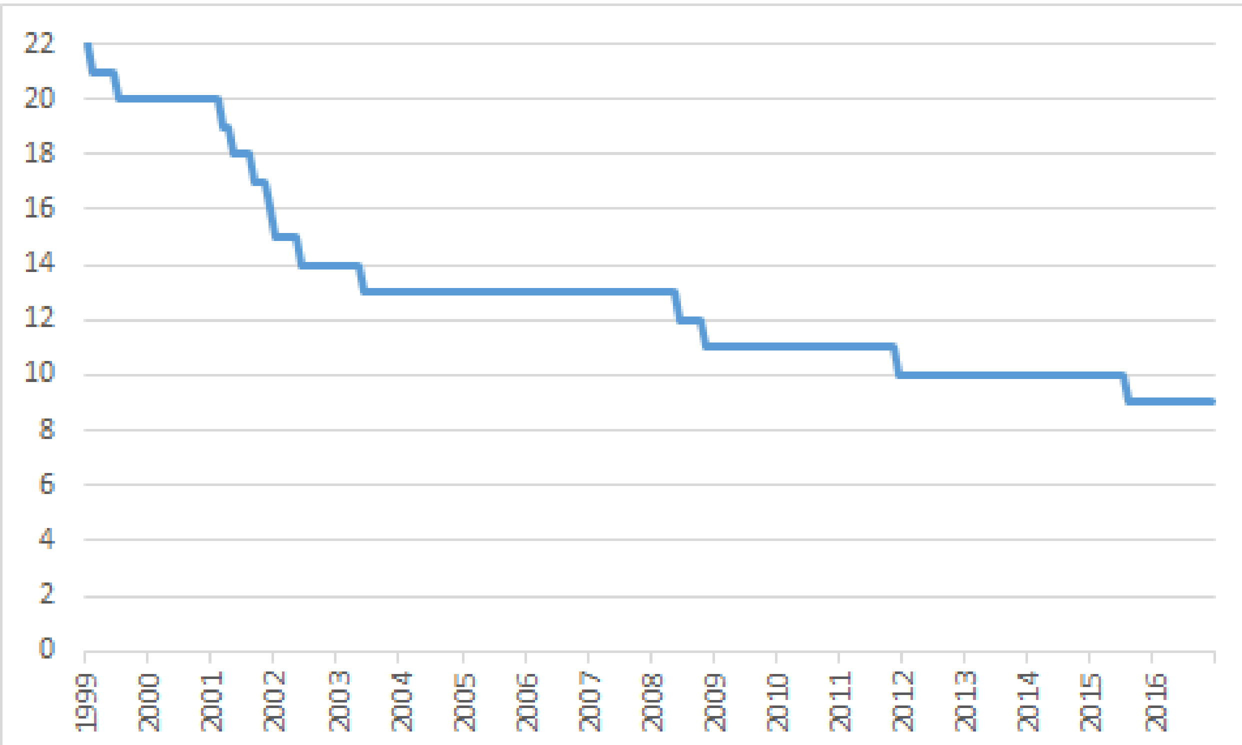


Note: Figure 1.4 reports the share of funds by country from 1999 to 2016 and the total TNA (in €Bn) of each of our funds' domicile country weighted by the number of funds available in this particular country.

A.2: Country-based analysis

The previous subsection discusses country-based categories. In this section, we explore the key features of the EEMF industry by considering style-based classifications. As presented in Section 3, our methodology enables the recovery of the number and composition of clusters over time. Recall that mutual funds included in the same cluster exhibit close dependence in their returns. One way to interpret these clusters is to consider that they group mutual funds that follow the same style and, as such, those set in competition. Figure 1.5 displays the number of style-based clusters from 1999 to 2016. On average, the industry can be divided into 13 styles. After a sharp decrease around the dotcom bubble, their number remains relatively stable over time, with slight decreases in 2008, 2011 and 2016. Our result points to an interesting feature. While more products have been made available in the industry EFAMA (2017), we find that it eventually goes along with more similarity in their returns. This finding should be further investigated with the help of portfolio holding data, for instance. This direction of research is, however, left for future investigations.

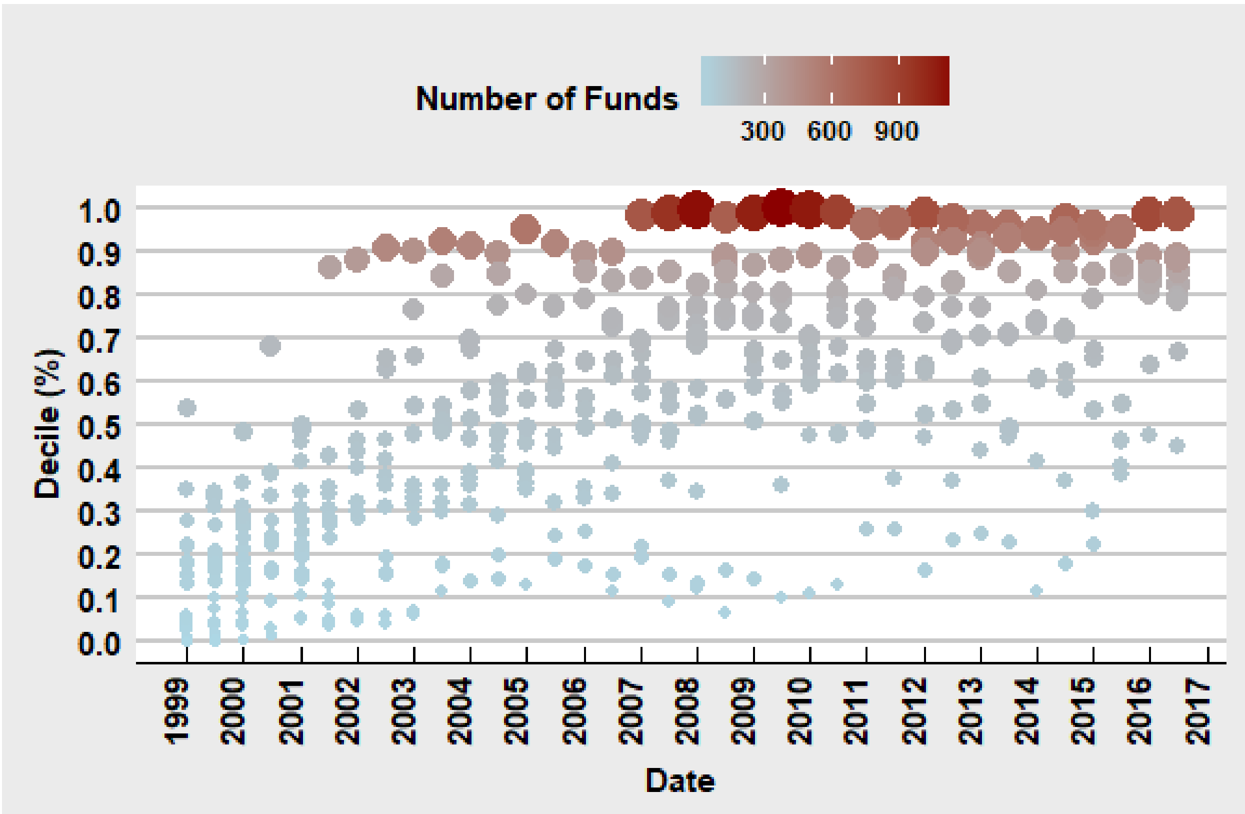
Figure 1.5: Time-varying cluster number



Note: Figure 1.5 reports the optimal number of clusters computed by the adaptive approach from 1999 to 2016.

Figure 1.6 reports the size of the style-based clusters over time. The y axis displays the decile to which each cluster is attached. Each cluster is represented by a circle whose size and color reflect the number of mutual funds included. Large clusters appear in 2001 and their number continues to increase. The largest structures are concentrated between 2007 and 2011. For instance, we count 1,170 mutual funds in a single cluster in March 2009. After 2012, we count a maximum of 966 institutions in a cluster. However, we observe more clusters in the last decile, supporting increased industry concentration.

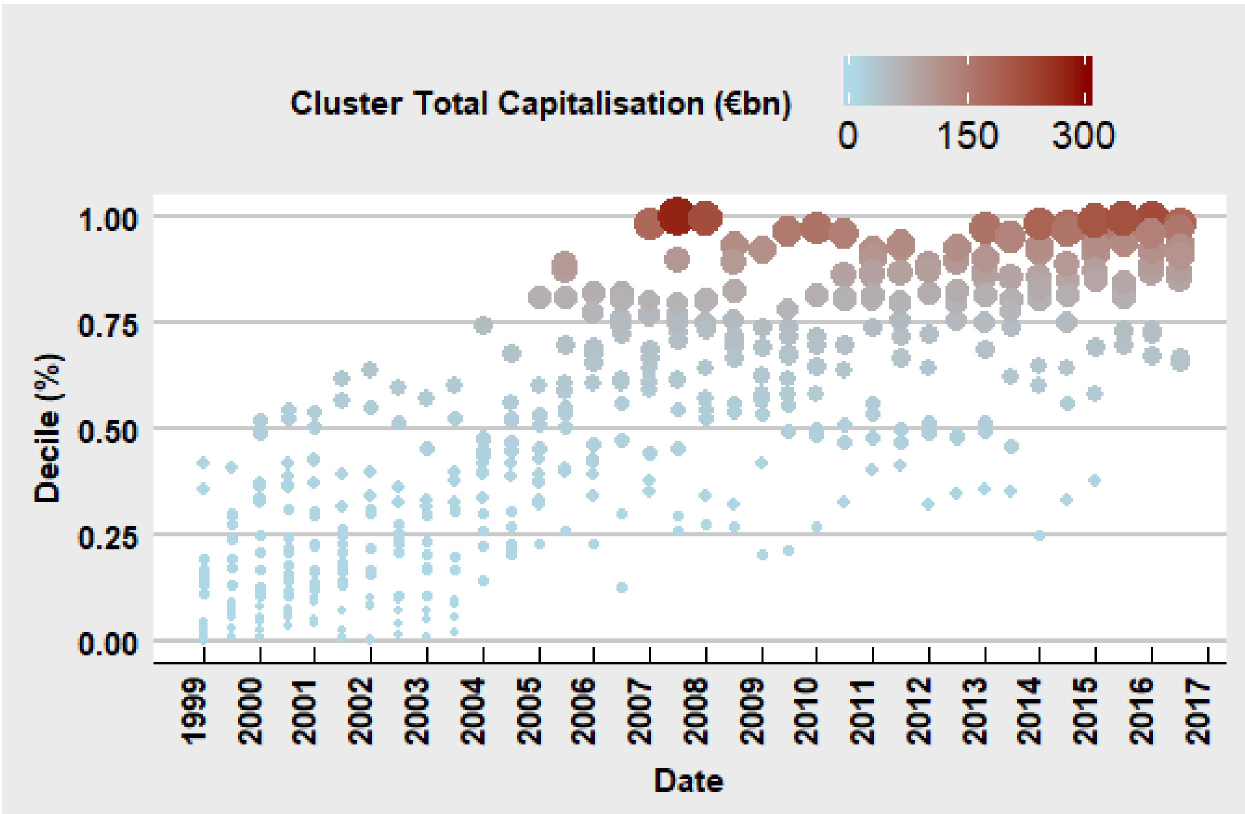
Figure 1.6: Size of style-based cluster



Note: Figure 1.6 reports the total number of funds (reported in deciles) by cluster from 1999 to 2016.

Figure 1.7 indicates the cluster size considering total AUM per cluster instead of the number of institutions. Two periods emerge. First, we observe strong concentration in the industry between 2007 and 2008. Next, the market becomes more fragmented before again displaying signs of increased concentration in 2011 and 2015. Larger clusters are visible in 2007–2008 and 2015–2016. Overall, we consistently find periods of strong integration that could go along with increased competition in 2001–2003, 2007–2008, and 2015–2016, characterized by a lower number of styles, more populated clusters, and larger sizes, along with phases of fragmentation in which we observe a stable number of smaller clusters, measured by the number of funds and smaller funds, as in 2005–2007 and 2011–2013.

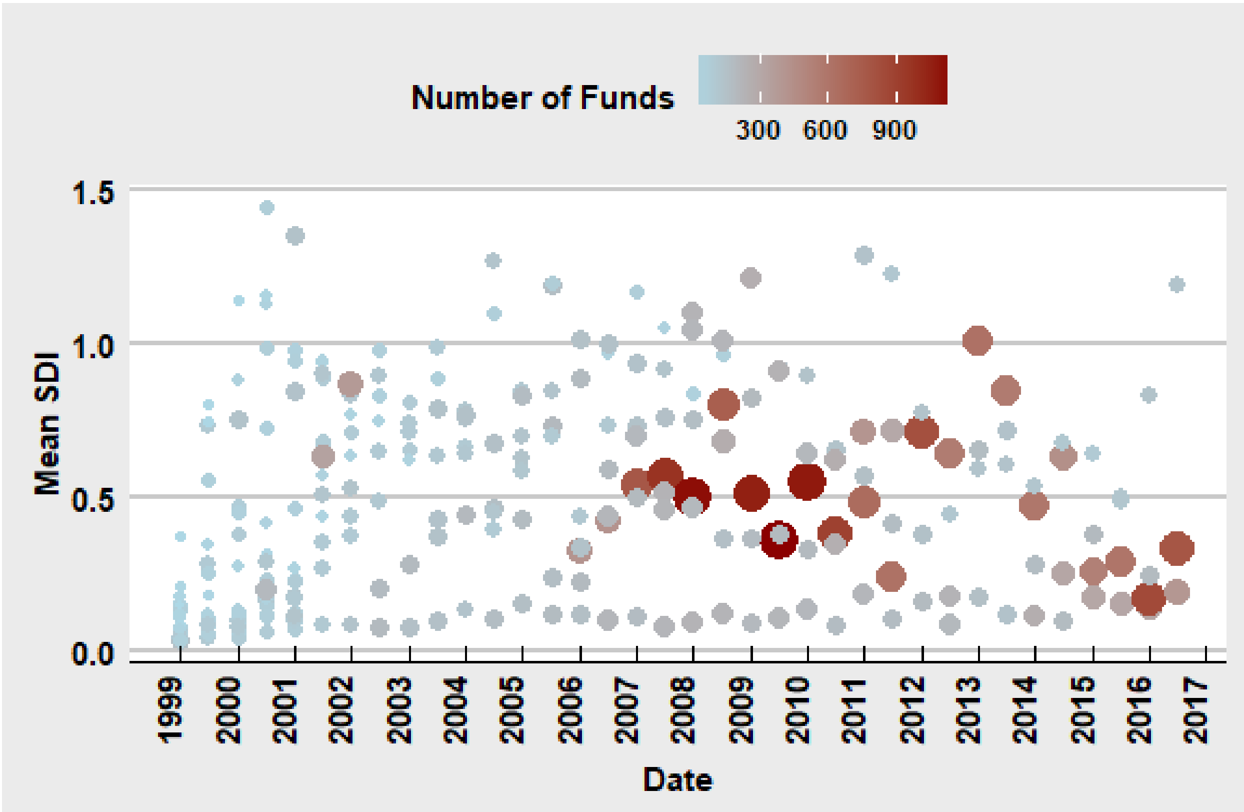
Figure 1.7: Total TNA of style-based cluster



Note: Figure 1.7 reports the average funds' TNA (reported in deciles) by cluster from 1999 to 2016.

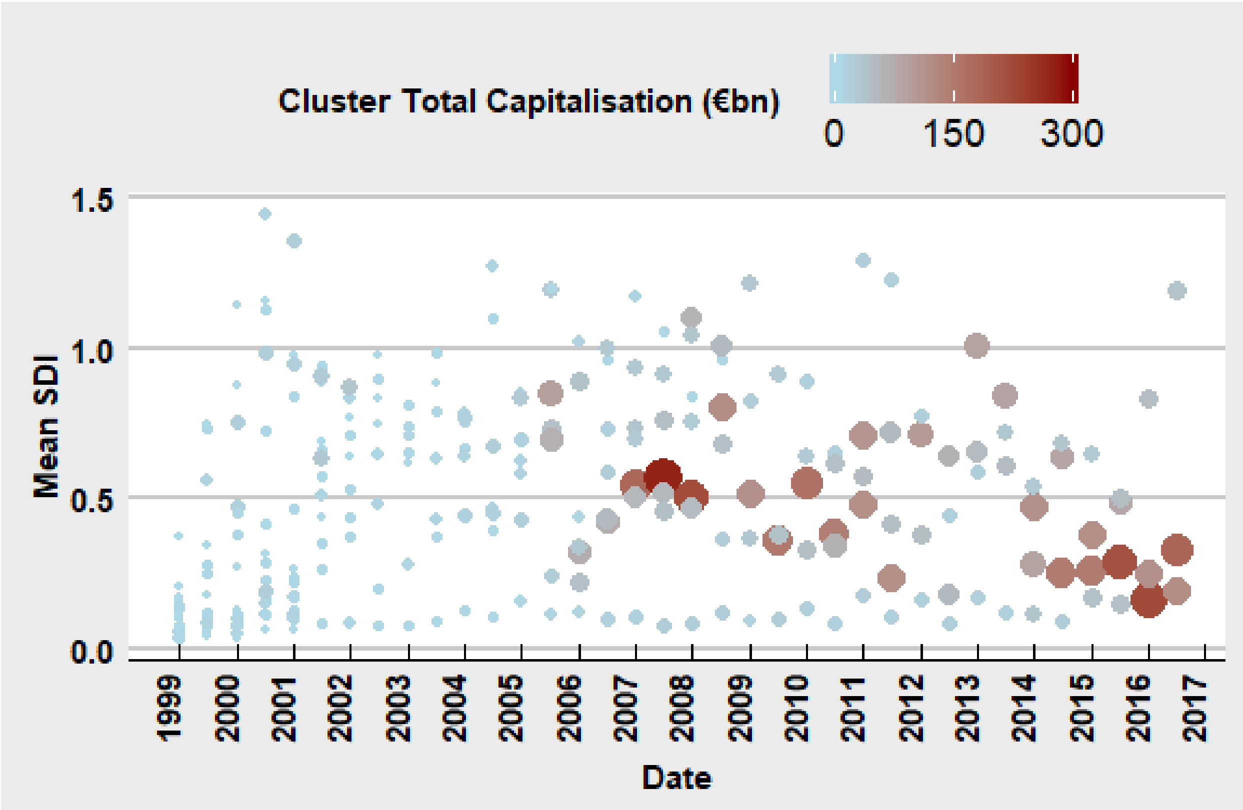
Next, we analyze the distinctiveness strategy of funds. Figures 1.8 and 1.9 display (i) the size of clusters as measured alternatively by the number of funds and AUM, along with (ii) the average SDI per cluster. Clusters with the highest average SDI are displayed in the upper part of the figure. By contrast, the most homogeneous clusters are displayed in the lower part of the figure. From the two figures, we do not observe a clear correlation between the level of distinctiveness and cluster size. For instance, in 2016, a large cluster is highly homogeneous, whereas in 2013, a fairly large cluster is heterogeneous. This bivariate analysis result tends to suggest that distinctive strategies are not more likely in a more populated and competitive environment.

Figure 1.8: Cluster mean SDI vs. number of funds



Note: Figure 1.8 reports the average funds' SDI (reported in deciles) and number of funds by cluster from 1999 to 2016.

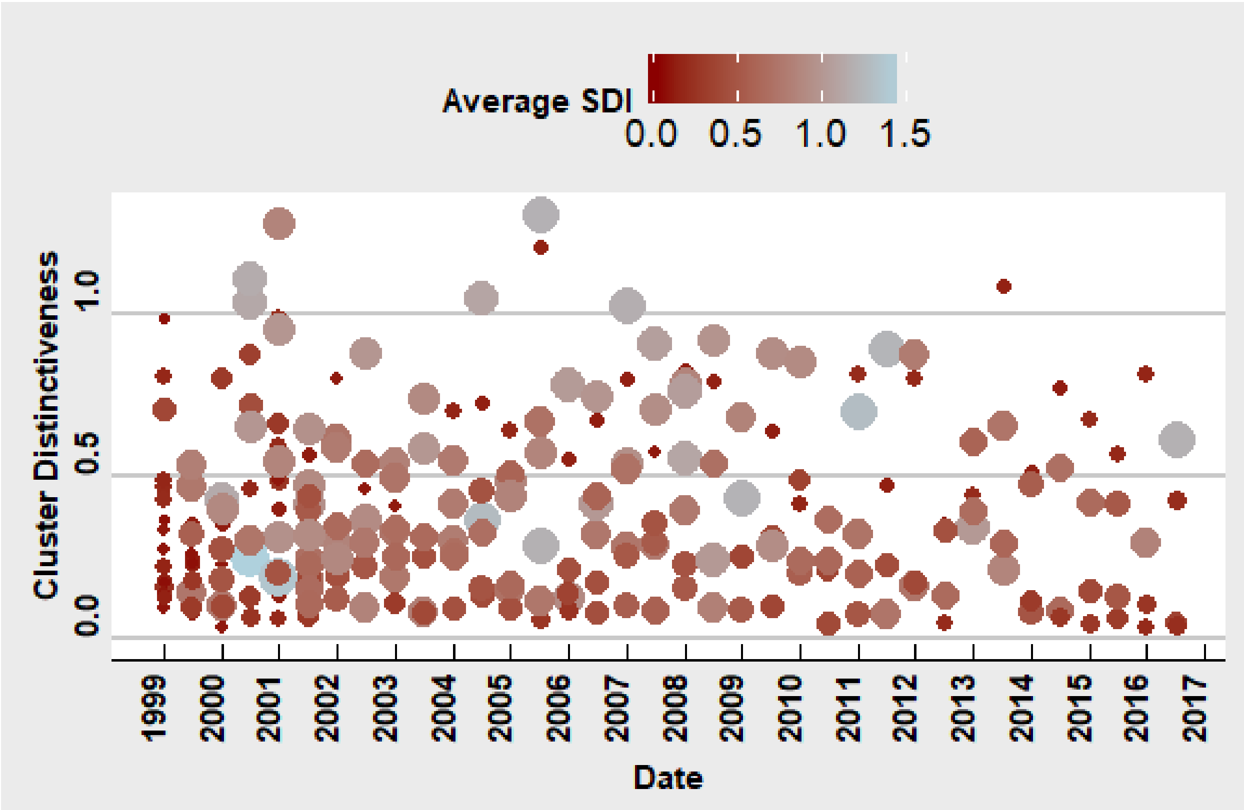
Figure 1.9: Cluster mean SDI vs. total capitalization



Note: Figure 1.9 reports the average funds' SDI (reported in deciles) and TNA by cluster from 1999 to 2016.

In the next figure, we compute the distinctiveness of the clusters themselves. To this end, we compare the average return of the entire industry with the average return of the cluster. Figure 1.10 shows the results. More distinct clusters are reported on the upper part of the figure. Clusters more in line with the rest of the industry are displayed in the lower part of the figure. The size of the circle depicts the average distinctiveness of the components of the cluster. Here, too, we cannot isolate specific regularities. For instance, in 2005, the two most distinct clusters are highly homogeneous and heterogeneous.

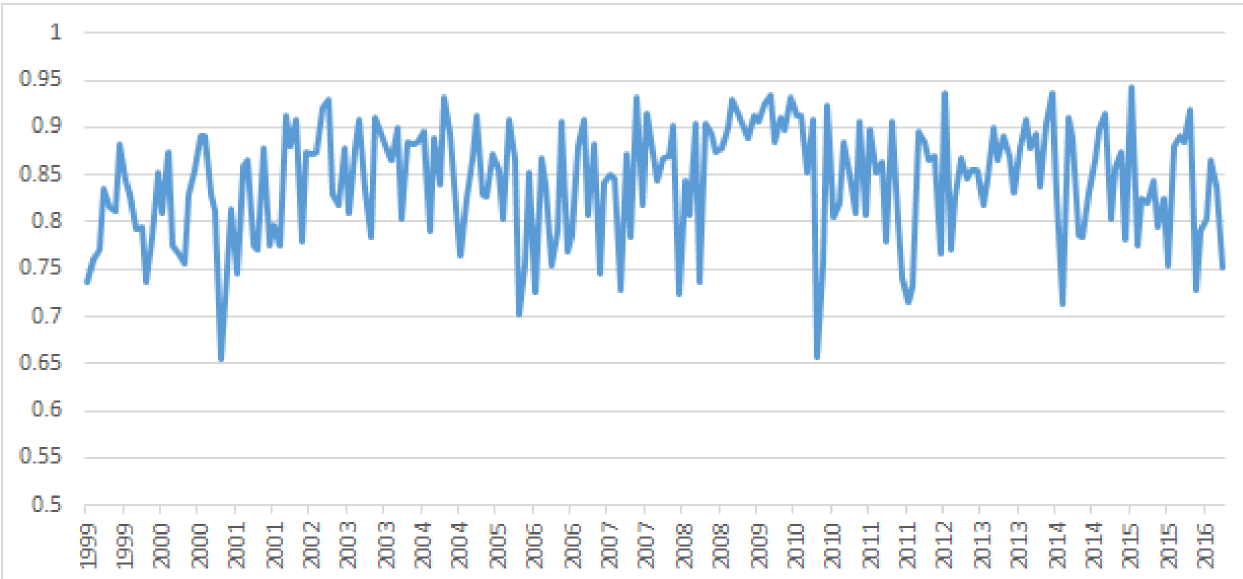
Figure 1.10: Cluster mean SDI vs. Cluster SDI



Note: Figure 1.10 reports the average Cluster-SDI (reported in deciles) and funds' average SDI by cluster from 1999 to 2016.

A final aspect that we discuss in this section concerns the dynamic nature of our system and, more specifically, changes in the composition of clusters and the SDI measure. An implicit assumption in the analysis is that funds select their style in the first place. Then, they proceed to make "marginal" adjustments in strategy to distinguish themselves from competing peers. In accordance, we should observe a high level of inertia in the composition of clusters and greater variability in the distinctiveness measure. Note, however, the two are not directly comparable. Nevertheless, we can show the persistence of the mutual fund cluster and whether such persistence has evolved over time. On average, a fund in two consecutive periods remains in the same cluster with an 85% probability. That is, 85% of funds at time t remain in the same cluster at time t+1. This finding supports the principle of inertia previously discussed. If we observe the evolution of this indicator, we see that persistence has exhibited strong consistency over time (Figure 1.11). Hence, the minimum match between two periods was achieved in 2000 (65% consistency).

Figure 1.11: Cluster stability across time



Note: Figure 1.11 reports the average stability in clusters' funds composition from one period to the next.